CONSUMER DEMAND FOR CONVENTIONAL FLUID MILK AND SELECTED DAIRY ALTERNATIVE BEVERAGES IN THE UNITED STATES

A Thesis

by

ALICIA MICHELLE COPELAND

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Senarath Dharmasena
Committee Members,	Oral Capps Jr.
	Ramkumar Janakiraman
Head of Department,	C. Parr Rosson III

August 2016

Major Subject: Agricultural Economics

ABSTRACT

Production and consumption of dairy alternative beverages in the United States has been on the rise as per capita consumption of fluid milk continues to fall. Almond milk is the fastest growing category in the U.S. dairy alternative marketplace, while soymilk popularity continues to decline. Many other dairy alternative beverages are also entering the marketplace, with the introduction of rice, hemp, hazelnut, coconut, and cashew milk in recent years. In order to determine the economic and demographic profiles of these dairy alternative beverage consumers, this study uses household-level data from the 2011 Nielsen Homescan panel which records household purchases from retail outlets for at home consumption. Utilizing the tobit econometric procedure, the conditional and unconditional own-price, cross-price and income elasticities for soymilk, almond milk, conventional white milk, and lactose-free milk were estimated.

Income, age, employment status, education level, race, ethnicity, region and presence of children are significant drivers affecting the demand for these dairy and dairy-alternative beverages. The conditional own-price elasticity of demand for almond milk, soymilk, lactose free milk and conventional fluid milk was estimated to be -0.55, -0.67, -0.49, and -0.69 respectively.

The elasticity estimates along with the demographic information can be used by beverage manufacturers and marketers to position dairy alternative beverages in the conventional dairy marketplace. Policy makers can also use this information to design appropriate policies for the U.S. dairy and dairy alternative beverage industries.

Future research in this area could include using the elasticity estimates to simulate the effects of dairy farmer welfare, shedding light on how the rising competition from dairy alternative beverages is impacting the dairy industry in the United States.

DEDICATION

I would like to dedicate this work to my mother and father. Thank you both so much for your endless love and support. I sincerely appreciate all you have done to get me to this point. I will forever look up to both of you and aspire to be like you in all aspects of my life.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Dharmasena and my committee members, Dr. Capps and Dr. Janakiraman for their support and encouragement throughout this process. I would especially like to thank Dr. Dharmasena for challenging me to attend and present this research at different conferences.

Thanks also go to my friends and cohort for making my experience at Texas A&M one I will never forget.

Finally, I would like to thank Mark Senia for his help on gathering the data for this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

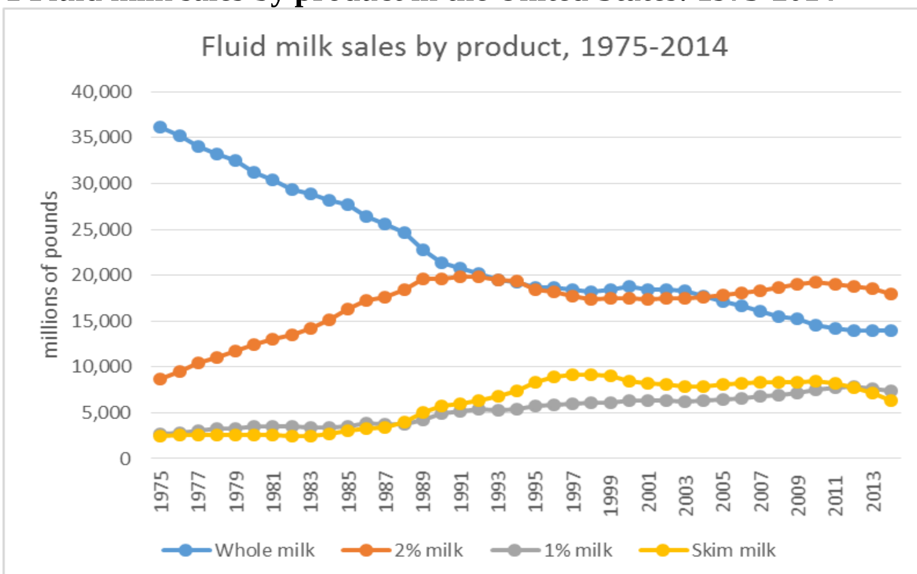
1. INTRODUCTION

Dairy products have always been staples in the American diet. Although over the past few years, changing consumer preferences and health concerns have forced dairy farmers and manufacturers to change production practices to produce new products to address these issues. Such innovations include the introduction of probiotics to conventional milk, new flavors, promotional incentives, and the addition of many vitamins, minerals, proteins, antioxidants, etc. (McCormack, 2016).

Even though the dairy industry strives to create and sell milk and processed dairy products that appeal to many consumers, per capita consumption of milk has been declining over the past 25 years. In fact, per capita consumption of milk has declined from 247 pounds per year in 1975 to 159 pounds per year in 2014 (USDA-ERS, 2015). This decline in demand for conventional dairy milk could be due to Americans becoming increasingly concerned with the fat content and additives (growth hormones and antibiotics) in their milk, which give conventional white milk the perception of being unhealthy.

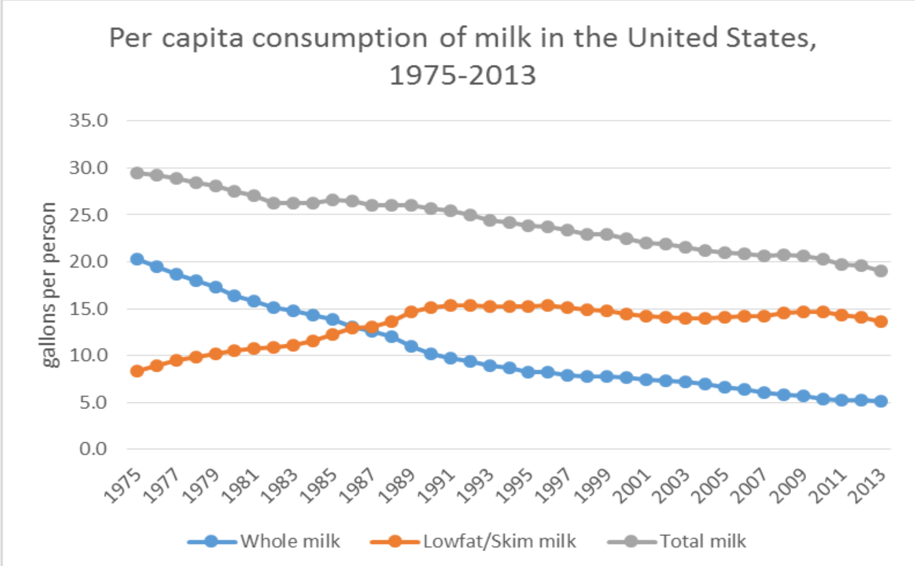
As seen in Figure 1, whole milk sales have been declining since 1975, with only 13,984 millions of pounds sold in 2014. Conversely, sales of milk products with lower fat content (2%, 1% and skim milk) have been increasing over the same period. This growth in the demand for low-fat milk, however, has somewhat leveled off after 2005, with a decreasing trend starting around 2011. Figure 2 also shows how per capita consumption of milk has fallen from nearly 30 gallons of milk per person in 1975 to below 20 gallons per person in 2014. More specifically, whole milk consumption has dropped from 20.3 gallons per person per year in 1975 to 5.2 gallons per person per year in 2013. Conversely, low-fat and fat-free milk consumption has increased from 8.4 gallons per person per year in 1975 to 19.0 gallons per person per year in 2013.

Figure 1 Fluid milk sales by product in the United States: 1975-2014



Source: USDA, Economic Research Service (2015)

Figure 2 Per capita consumption of milk in the United States: 1975-2013



Source: USDA, Economic Research Service (2015)

At the same time, sales of lactose-free dairy products in the United States have been increasing. In 2012, sales of lactose-free dairy products exceeded $700 million, and it is estimated that sales in this beverage category will continue to increase, amounting to $881.1 million in 2017 (Statista, 2016). With over 75% of the world's population being lactose intolerant, including 25% of people in the United States, it is no wonder that this beverage category continues to grow especially among certain ethnic groups. Approximately 90% of Asian Americans, 74% of Native Americans, 70% of African Americans, and 53% of Mexican Americans are lactose intolerant (Physicians Committee).

When comparing conventional white milk and dairy alternative beverages, consumers generally perceive dairy alternative beverages as being healthier. This perception could be because dairy alternative beverages have fewer calories per serving.

It could also be because these plant-based dairy alternatives do not come from animals treated with rBST, a bovine growth hormone. These plant-based milk products also appeal to Americans who are on restricted diets, such as veganism, gluten-free, and lactose intolerance. Many different types of plant-based dairy alternatives are available in a growing array of formulations. The most recognized dairy alternative beverages include almond milk, soymilk, rice milk, and coconut milk. New alternatives continue to become available, as hemp milk, oat milk, hazelnut milk, flax milk, sunflower milk, and multigrain milk are now a small part of the market. Growing health concerns of consumers combined with the availability of these dairy alternative beverages, which address these concerns, and allows Americans to still consume “milk” regularly, has led to constant growth for the producers of these plant-based milk beverages (Neville, 2015).

Soymilk was the original market leader of dairy alternative beverages, with soy being the primary or secondary ingredient in 78 percent of market launches in 2012 (Innova Market Insights, 2013). Almond milk demand has surged since its entry into the market, with an estimated 65.5% of total dairy alternative product demand. Soymilk currently represents only 25% of total dairy alternative product demand (Neville, 2015). Growth in dairy alternative beverages have been attributed to consumer perceptions of health-related claims, extended shelf life, appealing packaging, varying levels of sweetness, and the availability of flavors. The shift from soymilk to almond milk could be due to the fact that almond milk has no saturated fat, fewer calories than soymilk, and is rich in Vitamin E.

Sales of dairy alternative beverages in the United States reached nearly $2 billion in 2013, driven up largely because of the popularity of almond milk (Van Allen, 2014). The dairy alternative beverage industry in the United States is dominated by WhiteWave Foods (57.1% market share) who produces and markets almond milk and soymilk under the brand name Silk. Currently, Silk has 60% of the market for plant-based beverages with Silk Pure Almond being the main driver in the company's growth since 2010. Blue Diamond Growers is another industry leader with 21.2% market share. They produce and market almond milk under the brand name Almond Breeze. It is estimated that Almond Breeze holds 35% share of the almond milk market, with sales growing an average of 52.9% annually from 2010-2015 (Neville, 2015).

This increasing demand for dairy alternative beverages and declining demand for dairy milk in the United States could negatively affect dairy producers in terms of low prices for fluid milk as well as reduced farm income/welfare. Therefore, it is of interest for dairy producers in the United States to know the competitiveness of dairy alternatives in the dairy marketplace and their implications on dairy prices and farm income/welfare.

Furthermore, the producers and marketers of dairy alternative beverages in the United States would be interested in knowing the competitors of dairy alternative beverages in the marketplace. Such information would be useful for strategic decision-making, such as how to position these products in the marketplace, by U.S. dairy producers and cooperatives, as well as dairy alternative beverage producers and marketers.

The general objective of the research reported in this thesis is to evaluate the effect of dairy alternative beverages on the dairy market, specifically centering attention on almond milk as a dairy alternative beverage. The specific research objectives are to: (1) Estimate own-price, cross-price, and income elasticities of demand for almond milk, soy milk, conventional white milk, and lactose-free milk in the United States; (2) Estimate the economic and demographic profiles of demand for dairy alternative beverage consumer in the United States, more specifically centering attention on almond milk and soymilk.

This paper is organized as follows. Section 2 presents a discussion of literature on dairy product and dairy alternative beverage demand. The theoretical model is discussed in Section 3, followed by a description of the empirical model in Section 4. The data and variables used in this study are detailed in Section 5. Section 6 is dedicated to discussing the empirical results. Finally, concluding remarks, implications, and suggestions for future work are found in Section 7. Derivations and the SAS Code used to complete this research can be found in the Appendix.

2. LITERATURE REVIEW

Conventional Milk Market Research

Previous work on consumer demand for dairy milk and dairy alternative beverages has given us some important insights into these markets. Gould, Cox and Perali (1990) examined how the changing demographic structure of the U.S. impacted demand for fluid milk products. The Almost Ideal Demand System (AIDS) model which was formulated by Deaton and Muellbauer (1980) was used in this study. Unlike many previous studies which were cross-sectionally based analyses that considered only food-at-home consumption, this study used time-series data from the period of 1955 to 1985 and included food-away-from-home consumption as well. The five products included in this analysis were whole milk, low-fat milk, fruit juices, other nonalcoholic beverages, and other food. The other beverages category is composed of coffee, tea and nonalcoholic carbonated beverages. The demographic variables included in the demand model were different age group proportions, the proportion of the population that is non-white, and the median number of years of schooling completed for those over 25 years of age. All beverages were found to have inelastic own-price elasticities at -0.324, -0.437, -0.37, -0.193 and -0.40 for whole milk, low-fat milk, juices, other beverages and other food respectively.

Kaiser and Reberte (1996) noted that a common characteristic of all previous studies of generic fluid milk advertising was that they aggregated fluid milk products into a single product. In order to address this issue, they disaggregated fluid milk products into whole, low-fat and skim milk categories in order determine if individuals who purchased the various types responded differently to the existing advertising strategy for fluid milk. Kaiser and Reberte (1996) used a log-log model to estimate the demand for each of the fluid milk products. The explanatory variables included were the retail price of each of the three fluid milk products; retail price of orange juice; disposable per capita income; a variable measuring consumer concerns about dietary fat; seasonal quarterly dummy variables; and per capita advertising expenditures. The dependent variable was the per capita sales of fluid milk products in the New York City metropolitan area. Separate demand functions were estimated for each milk category using monthly time-series data from the New York-New Jersey federal milk marketing order from 1986 to 1992. All estimated own-price elasticities were inelastic. The own-price elasticities were: -0.003 for whole milk, -0.14 for low-fat milk and -0.30 for skim milk. The fat concern variable was only found to be significant in the demand equation for whole milk. Finally the results showed that advertising for generic fluid milk positively impacted the sales of all three milk categories in the New York metropolitan area.

Gould (1996) estimated demand for milk within a system-wide framework. Gould utilized Nielsen household panel data with over 4,300 households to determine the own-price and cross-price elasticities of whole milk, skim (1%) milk, and 2% milk. This research found that the own-price elasticities were -0.803, -0.593 and -0.512 for whole milk, skim (1%) milk and 2% milk respectively. The study also found that the three milk types investigated were substitutes in consumption. This study is one of the few econometric studies involving fluid milk demand that incorporates the substitution possibilities across milk fat types and also incorporates the censored nature of the data set.

Davis, Dong, Blyaney, Yen and Stillman (2012) examined the impact of demographic variables, retail prices and total milk expenditure on flavored and non-flavored milk purchases. In their analysis, they used the 2007 household data from Nielsen Homescan. Unlike previous research, they estimated demand for conventional white milk and flavored milk using four different levels of fat content, whole, 2%, 1% and skim. They used a censored AIDS model to derive the elasticity estimates. The conditional own-price elasticities for the conventional white milk were estimated to be -0.67, -0.44, -0.71 and -0.80 for whole milk, 2% milk, 1% milk and skim milk respectively. Their results also showed that the presence of children positively influences retail purchases of whole milk, 2% milk and 1% milk.

Nonconventional Milk and Dairy Alternatives Market Research

Glaser and Thompson (2000) examined retail sales for organic and conventional milk. During the years prior to this study, organic milk and many other organic products were entering mainstream supermarkets. Glaser and Thompson (2000) obtained data from three sources, Spince Information Services (SPINS), Nielsen scanner data and Information Resources, Inc. (IRI). The data was separated into three milk types, organic, private label and branded. Each milk type was then broken into four categories (whole, 2%, 1% and nonfat/skim) based on fat content. The nonlinear almost-ideal demand system (AIDS) was used to estimate four demand systems. Branded milk and private label milk were both found to be own-price inelastic. Organic milk, however was found to be highly own-price elastic. Branded and organic milks were found to be substitutes in consumption.

Alviola and Capps (2010) estimated sociodemographic profiles of conventional and organic dairy milk consumers in the United States. The study used the 2004 Nielsen panel data, which consists of over 38,000 households, to identity the drivers of the demand for conventional and organic milk in the United States. In particular, they wanted to investigate the own-price effects, the cross-price effects, the income effects, and the effects of the sociodemographic characteristics on household decisions to purchase organic or conventional milk. After the decision to purchase organic milk or conventional milk had been made, the researchers then focused on the factors that affected how much of each type of milk was purchased. In order to complete this work, Alviola and Capps (2010) employed the Heckman two-step procedure. They also addressed issues of price endogeneity, since the prices were derived as the ratio of total expenditures to total quantity purchased, by conducting Hausman tests. The results from this study indicated that organic milk and conventional milk are substitutes. The elasticities also indicate that demand for organic milk is more sensitive to changes in price of conventional milk, but that the demand for conventional milk is not very sensitive to changes in the price of organic milk. Also, household size, number of children, employment status/education of household head, race, ethnicity, and region had a significant impact on the likelihood of a household to purchase organic milk.

Lopez and Lopez (2009) analyzed consumer choices price competition in the differentiated fluid milk market. This research applies the model developed by Berry, Levinson, and Pakes (1995) known as the BLP model, to a sample of fluid milk products in the Boston area. The milk sales information came from Information Resources Incorporated (IRI) database and consists of milk sold by the four leading supermarket chains in Bristol, Essex, Middlesex, Norfolk, Plymouth, Suffolk, and Worcester counties. Product characteristics included brand name, fat content, lactose content, and organic milk. The brand names included private label, Garelick, Hood, Organic Cow of Vermont, Morningstar, and McNeil. The fat contents included in this study were 0%, 1%, 2% and whole milk, which is 3.25%. The own-price elasticities for all milk categories were negative, with private label milk being the least price sensitive. The specialty milks, which are more expensive, had higher own-price elasticities. The cross-price elasticities were more intense within milk categories and greater among products with the same fat content. The own and cross-price elasticities also showed that when the price of lactose-free brands increased, people stop buying lactose-free milk. There is evidence of significant substitution across lactose-free brands and limited substitution to other milk types.

Zheng (2011) compared consumer demand for soymilk in the United States and China. Specifically, in the United States, Zheng (2011) identified the consumption behavior of soymilk consumers and assessed their preferences for organic soymilk. Using an on-line survey delivered to random respondents via a market research company, a choice experiment was conducted to determine the willingness to pay for soymilk attributes. Mixed logit models were then used to analyze the choice experiment response. The results showed that 70% of respondents said Silk was their first choice when purchasing soymilk, which confirms Silks dominant market position. In addition, 61.71% of respondents perceived a difference in the “taste or flavor” of store-brand and national brands. About 66% of respondents shopped for soymilk more than once a month. The respondents indicated that the three most important attributes in soymilk were: taste, minimum use of preservatives and low risk of food-borne illness. The least important attribute was lactose or casein free, indicating that consumers wanted to drink soymilk for reasons other than an allergy.

Choi and Wohlgenant (2012) analyzed demand for fluid milk using Nielsen Homescan panel data. The fluid milk products purchased were differentiated by fat content, flavor and organic claim. Hausman’s three stage demand systems approach was use to estimate the demand for the fluid milk products. All milk products had negative own price elasticities except organic-flavored soy/lactose free milk. Own price elasticities for conventional milk and organic milk ranged from -1.36 to -2.71 and -1.00 to -7.34, respectively. The cross price elasticities estimated showed different substitution patterns than previous studies. For instance, the cross-price elasticities between organic and conventional milk, with the same fat content and flavor, did not confirm substitutability between the goods. The cross-price elasticities also indicated that only milk products with similar fat content are substitutes for each other.

Dharmasena and Capps (2014) investigated U.S. consumer demand for dairy alternative beverages, more specifically soymilk. In this research, the conditional and unconditional factors that affect the volume of soymilk, white milk, and flavored milk purchased were identified. The study also determined the conditional and unconditional own-price, cross-price and income elasticities of demand for soymilk, white milk and flavored milk. Finally, this work provided retail-level pricing strategies for soymilk, white milk, and flavored milk in the marketplace. Dharmasena and Capps (2014) utilized the Tobit model due to the censored nature of the 2008 Nielsen Homescan data. Statistically significant determinants of demand for soymilk were age of household head, employment status of household head, education status of household head, region, race, Hispanic status, age and presence of children, and gender of household head. The results of this study showed that white milk and flavored milk were gross substitutes for soymilk. This research study also demonstrated that the conditional own-price elasticity of demand for soymilk was -0.30 meaning that consumers are insensitive to changes in its own price.

3. THEORETICAL MODEL

Data collected at the consumer level, such as individual or household level, often results in consumers or individuals not purchasing all or some of the products during the sampling period due to various reasons. The 2011 Nielsen Homescan panel data set used in this study contains households that did not purchase all of the products considered. In this data set, of the total number of households, some did not purchase soymilk, almond milk, conventional white milk, and/or lactose-free milk during the sampling period. The presence of these responses in the data creates a zero consumption level for that observation, meaning the dollar amount that household spent was recorded as a zero. This situation creates a *censored* sample of data.

Application of ordinary least squares (OLS) to estimate a regression with a censored dependent variable can result in biased estimates, even asymptotically (Kennedy, 2003). If all observations with zero purchases were removed and only non-zero purchase observations were used to estimate regression functions, there would be sample selection bias (Kennedy, 2003). Tobin (1958) and Heckman (1979) suggested alternative models to deal with sample selection bias in estimated regression models in the presence of censored data. Using Tobin's model for a dependent variable that takes on a zero value with positive probability but continuously distributed over positive values, both conditional and unconditional elasticity estimates pertaining to soymilk, almond milk, lactose-free milk, and conventional white milk can be obtained. Also, in order to provide insight on the probability of being above the limit (or probability of purchase) where the limit being zero in this analysis, and changes in the value of the dependent variable if it is already above the limit, McDonald and Moffitt's (1980) decomposition of the coefficient estimates of the Tobit model also can be calculated.

The stochastic model underlying the tobit model can be expressed as follows:

(1) $$y_i = \begin{cases} \boldsymbol{X}_i\boldsymbol{\beta} + u_{i,} & \boldsymbol{X}_i\boldsymbol{\beta} + \mu_i > 0 \\ 0, & \boldsymbol{X}_i\boldsymbol{\beta} + \mu_i \leq 0 \end{cases} \qquad \mu_i \sim N(0, \sigma^2)$$

where $i = 1,2,3, \ldots\ldots, N$, the number of observations. y_i is the censored dependent variable; $\boldsymbol{X}_i$ is the vector of explanatory variables; $\boldsymbol{\beta}$ is the vector of unknown parameters to be estimated, μ_i is the normally distributed error. Using the Tobit model gives rise to two expectations for the dependent variable y. Equation (2) expresses the unconditional expected value (observations that are at and above the limit) for y_i and equation (3) expresses the corresponding conditional expected value (observations that are above the limit) for y_i. In both of these equations, the normalized index value z is shown as $z = \frac{X\beta}{\sigma}$, where σ is the estimated standard error of the tobit regression. Further, $F(z)$ represents the cumulative distribution function (CDF) and $f(z)$ represents the probability density function (pdf) associated with the normalized index value, z.

(2) Unconditional Expected Value: $E(y) = \boldsymbol{X\beta}F(z) + \sigma f(z)$

(3) Conditional Expected Value: $E(y^*) = \boldsymbol{X\beta} + \sigma \frac{f(z)}{F(z)}$

The coefficients in the Tobit model represent the effect of an independent variable on the latent dependent variable. Therefore, the marginal effects must be calculated. As with the expected values, there are two types of marginal effects. The first is the unconditional marginal effect for consumption of a beverage by all households, whether they purchased the beverage or not. The second type of marginal effect is the conditional marginal effect, which reflects the marginal effect for consumption of a beverage only by households that bought the beverage. Equation (4) represents the unconditional marginal effect, and equation (5) shows the conditional marginal effect.

(4) Unconditional Marginal Effect: $\frac{\partial E(y)}{\partial \boldsymbol{X}} = \boldsymbol{\beta} F(z).$

(5) Conditional Marginal Effect: $\frac{\partial E(y^*)}{\partial \boldsymbol{X}} = \boldsymbol{\beta}(1 - z\frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2}).$

In order to relate the conditional expectation and the unconditional expectation, the McDonald and Moffitt (1980) decomposition be shown in equation (6) can be used. As equation (6) shows, the McDonald and Moffitt (1980) decomposition equates the total change in the unconditional expected value of the dependent variable, *y* to the sum of the change in the conditional expected value of *y* being above the limit, weighted by the probability of being above the limit and the change in probability of being above the limit weighted by the conditional expected value of *y* being above the limit. It should also be noted that, the expression in equation (6) equals $\boldsymbol{\beta} F(z)$ as shown in equation (4).

(6) $\frac{\partial E(y)}{\partial X} = F(z)\left(\frac{\partial E y^*}{\partial X}\right) + E(y^*)\left(\frac{\partial F(z)}{\partial X}\right)$

4. EMPIRICAL MODEL

In this sample of data, the dependent variable takes on a wide range of values. For some significant fraction, quantity purchased was zero for each beverage. For those households in which the quantity purchased was zero, there was also zero expenditures since they did not purchase anything. Consequently, there is not an observation of any unit value or price for those households. For households that did purchase a beverage, both the quantity purchased and total expenditure were recorded. Since the Nielsen data does not record a price, a unit value is used as a proxy for price. This unit value is calculated by dividing total expenditure by quantity for each beverage category. Since the price of each beverage category is the used as an explanatory variable in the model, prices need to be imputed for observations where no price was observed. To accomplish this imputation, an auxiliary regression, where observed prices for each beverage were regressed on household income, household size, and region where the household is located, was done. This is a standard procedure used in the price imputation literature, for example see Capps, et al, (1994); Alviola and Capps (2010); Kyureghian, Nayga and Capps (2011); and Dharmasena and Capps (2014).

$$P_{i,observed} = a_1 + (a_2 \times HH_{i,income}) + \left(a_3 \times HH_{i,size}\right) + \left(a_4 \times HH_{i,region}\right) + \mu_i \tag{7}$$

Where *i=1,2,3,….n*, number of households.

The three variables used in the auxiliary regression all address different issues that are often reflected in the price of a good. The income variable reflects the variability of demand for different quality of beverages. Household size reflects various socio-demographic conditions. The region in which the household is located reflects how prices differ based on location. Once calculated, forecasted prices were used as observations in which no price was observed. The prices for each beverage (soymilk, almond milk, lactose-free milk and conventional white milk) were then used as explanatory variables to estimate each tobit model pertaining to soyilk, almond milk, lactose-free milk and conventional white milk consumption.

Summary statistics for the observed and imputed prices for each beverage category are shown in Table 1. The means and standard deviations are consistent for with-in sample estimates as well as out-of-sample price imputations. The mean of the observed price of almond milk is $0.0530/ounce which is consistent with the mean of the imputed price of $0.0531/ounce. The mean of the observed price for soymilk is $0.0547/ounce which is consistent with the mean of the imputed price of $0.0548/ounce. The mean of the observed price of white milk is $0.0300/ounce which is very consistent with the mean of the imputed price of $0.0301/ounce. The mean of the observed price for lactose-free milk is $0.0565/ounce which is consistent with the mean of the imputed price of $0.0561/ounce.

Table 1 Summary Statistics of Observed and Imputed Prices of Almond Milk, Soymilk, Conventional White Milk, and Lactose-free Milk

	Observed Price (U.S. dollars per ounce)		Imputed Price (U.S. dollars per ounce)	
	Mean	Standard Deviation	Mean	Standard Deviation
Almond Milk	0.0530	0.0130	0.0531	0.0020
Soymilk	0.0547	0.0167	0.0548	0.0017
Conventional White Milk	0.0300	0.0125	0.0301	0.0121
Lactose-free Milk	0.0565	0.0113	0.0561	0.0045

Source: Calculated by the author

Pearson correlation coefficients for the four beverage prices are displayed in Table 2. The correlations among all beverages are statistically significant under a p-value of 0.05. The correlation between almond milk price and soymilk price is equal to 0.191. The correlation between almond milk price and white milk price is 0.019. The correlation between almond milk price and lactose-free milk price is 0.250. The correlation between soymilk price and white milk price is 0.033. The correlation between soymilk price and lactose-free milk price is 0.162. Finally, the correlation between white milk price and lactose-free milk price is 0.087. However, it must be noted that the correlation between the dairy alternative beverages (almond milk and soymilk) as well as between lactose-free milk and conventional white milk are positive, yet small.

Table 2 Correlation Test for Beverage Prices

	Almond Milk Price	Soymilk Price	Conventional White Milk Price	Lactose-free Milk Price
Almond Milk Price	1	**0.191** (<.0001)	**0.019** (<.0001)	**0.250** (<.0001)
Soymilk Price		1	**0.033** (<.0001)	**0.162** (<.0001)
Conventional White Milk Price			1	**0.087** (<.0001)
Lactose-free Milk Price				1

Source: Computed by the author
Notes: Pearson correlation coefficients, N = 62092, Prob > |r| under H0: Rho=0

Pearson correlation coefficient was also calculated for the error terms to determine whether it was necessary to use a system of equations to estimate the model, rather than four separate single equation Tobit models. The results of this test are displayed in Table 3. The correlation among all error terms is statistically significant under a p-value of 0.05, yet again small, not warranting the estimation of a system of equations. The correlation between the almond milk error term and the soymilk error term is 0.410. The correlation between the almond milk error term and the white milk error term is -0.026. The correlation between the almond milk error term and the lactose-free milk error term is 0.399. The soymilk error term and the white milk error term are negatively correlated with a correlation of -0.014. The soymilk error term and the lactose-free milk error term is 0.394. Finally, the white milk error term and the lactose-free milk error term are negatively correlated with a correlation of -0.033.

Table 3 Correlation of the Error Terms of the Respective Tobit Equations

	Almond Milk	Soymilk	Conventional White Milk	Lactose-free Milk
Almond Milk	1	**0.410** (<.0001)	**-0.026** (<.0001)	**0.399** (<.0001)
Soymilk		1	**-0.014** 0.0006	**0.394** (<.0001)
Conventional White Milk			1	**-0.033** (<.0001)
Lactose-free Milk				1

Source: Computed by author
Notes: Pearson correlation coefficients, N = 62,092, Prob > |r| under H0: Rho=0

Several hypotheses regarding purchases of almond milk, soymilk, conventional white milk, and lactose-free milk were tested. The specific hypotheses tested are as follows: (1) Conventional white milk and the dairy alternative beverages almond milk and soymilk are gross substitutes in consumption and will therefore have positive cross-price elasticities. (2) Conventional white milk and lactose-free milk are gross substitutes in consumption. (3) Households with higher levels of education will consume more dairy alternative beverages and less white milk. (4) Households where the household head is employed full time will consume less white milk. (5) Households that identify as non-white will consume more dairy alternative beverages and lactose-free milk than white households will. (6) Households located in the Pacific region of the United States will consume more dairy alternative beverages. (7) Households with children will consume more white milk than the households with no children.

A single-equation Tobit model for each beverage category (soymilk, almond milk, conventional white milk and lactose-free milk) was estimated. Several functional forms such as linear, quadratic and semi-log were investigated to find which model performs best based on the following criteria, model fit, significance of variables and loss metrics such as the Akaike Information Criteria (AIC), Schwarz Information Criteria (SIC) and Hannan-Quinn Information Criteria (HQC). A semi-log functional form outperformed the other functional forms and was thus used to estimate both conditional and unconditional marginal effects associated with each explanatory variable. The level of significance used in this study is 95% (*p*-value is 0.05). The demand functions for the four beverages are as follows:

(8) Soymilk demand model:

$$y_{\text{quantity, soy}} = \delta_1 + \delta_2 \text{logprice}_{\text{soymilk}} + \delta_3 \text{logprice}_{\text{almondmilk}} + \delta_4 \text{logprice}_{\text{lactosefreemilk}} + \delta_5 \text{logprice}_{\text{whitemilk}} + \delta_6 \text{logHHincome} + \delta_D \mathbf{X}_{\text{demographic variables}} + \mu_{\text{i, soy}}$$

(9) Almond milk demand model:

$$y_{\text{quantity, almond}} = \rho_1 + \rho_2 \text{logprice}_{\text{soymilk}} + \rho_3 \text{logprice}_{\text{almondmilk}} + \rho_4 \text{logprice}_{\text{lactosefreemilk}} + \rho_5 \text{logprice}_{\text{whitemilk}} + \rho_6 \text{logHHincome} + \rho_D \mathbf{X}_{\text{demographic variables}} + \mu_{\text{i, almond}}$$

(10) Lactose-free milk demand model:

$$y_{quantity,lactosefree} = \varphi_1 + \varphi_2 logprice_{soymilk} + \varphi_3 logprice_{almondmilk} + \varphi_4 logprice_{lactosefreemilk} + \varphi_5 logprice_{whitemilk} + \varphi_6 logHHincome + \varphi_D \boldsymbol{X}_{demographic\ variables} + \mu_{i,lactosefree}$$

(11) Conventional white milk demand model:

$$y_{quantity,whitemilk} = \omega_1 + \omega_2 logprice_{soymilk} + \omega_3 logprice_{almondmilk} + \omega_4 logprice_{lactosefreemilk} + \omega_5 logprice_{whitemilk} + \omega_6 logHHincome + \omega_D \boldsymbol{X}_{demographic\ variables} + \mu_{i,whitemilk}$$

Where $\boldsymbol{X}_{demographic\ variables}$= $X_{age\ of\ household\ head\ 25-29}$, $X_{age\ of\ household\ head\ 30-34}$, $X_{age\ of\ household\ head\ 35-44}$, $X_{age\ of\ household\ head\ 45-54}$, $X_{age\ of\ household\ head\ 55-64}$, $X_{age\ of\ household\ head\ 65\ or\ older}$, $X_{employment\ status\ part-time}$, $X_{employment\ status\ full-time}$, $X_{education\ high\ school}$, $X_{education\ undergraduate}$, $X_{education\ post-college}$, X_{Black}, X_{Asian}, X_{other}, $X_{HispanicOrign}$, $X_{New\ England}$, $X_{Middle\ Atlantic}$, $X_{East\ North\ Central}$, $X_{West\ North\ Central}$, $X_{South\ Atlantic}$, $X_{East\ South\ Central}$, $X_{West\ South\ Central}$, $X_{Moutain}$, $X_{children\ less\ than\ 6\ years}$, $X_{children\ 6-12\ years}$, $X_{children\ 13-17\ years}$, $X_{children\ under\ 6\ and\ 6-12\ years}$,

$X_{children\ 6-12\ and\ 13-17\ years}$, $X_{children\ under\ 6,\&\ 6-12,\&\ 13-17}$, $X_{female\ head\ only}$, $X_{male\ head\ only}$.

The following derivations and results are based off the semi-log functional form. The equation for the unconditional marginal effect for the price and the conditional marginal effect for the price variable are expressed in equation (12) and equation (13) respectively.

(12) Unconditional Marginal Effect: $\frac{\partial E(y)}{\partial p} = \frac{\beta}{P^U} F(z)$

where P^U is the unconditional average price (all of the observations) for each beverage considered.

(13) Conditional Marginal Effect: $\frac{\partial E(y^*)}{\partial p} = \frac{\beta}{P^C}(1 - z\frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2})$

where, p^C is the conditional average price of censored sample for each beverage considered.

The unconditional own-price, cross-price and income elasticities are represented by equations (14), (15) and (16) respectively,

(14) Unconditional Own-Price Elasticity: $\varepsilon_{ii}^U = \frac{\beta}{P_i^U} F(z) \frac{P_i^U}{Q_i^U}$

(15) Unconditional Cross Price Elasticity: $\varepsilon_{ij}^U = \frac{\beta}{P_j^U} F(z) \frac{P_j^U}{Q_i^U}$

(16) Unconditional Income Elasticity: $\varepsilon_I^U = \frac{\beta}{I_i^U} F(z) \frac{I_i^U}{Q_i^U}$

where, ε_{ii}^{U} is the unconditional own-price elasticity for *i;* ε_{ij}^{U} is the unconditional cross price elasticity which measures the change in quantity demanded for beverage *i* when there is a change in price of beverage *j*; ε_{I}^{U} is the unconditional income elasticity for the *i*th beverage.

The conditional own-price, cross-price and income elasticities are represented by equations (17), (18), (19) respectively,

(17) Conditional Own-Price Elasticity: $\varepsilon_{ii}^{C} = \frac{\beta}{P_i^C}(1 - z\frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2})\frac{P_i^C}{Q_i^C}$

(18) Conditional Cross Price Elasticity: $\varepsilon_{ij}^{C} = \frac{\beta}{P_j^C}(1 - z\frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2})\frac{P_j^C}{Q_i^C}$

(19) Conditional Income Elasticity: $\varepsilon_{I}^{C} = \frac{\beta}{I_i^C}(1 - z\frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2})\frac{I_i^C}{Q_i^C}$

where, ε_{ii}^{C} is the conditional own-price elasticity for *i;* ε_{ij}^{C} is the conditional cross-price elasticity which measures the change in quantity demanded for beverage *i* when there is a change in price of beverage *j*; ε_{I}^{C} is the conditional income elasticity for the *i*th beverage.

This research does not impose restrictions of demand theory, such as homogeneity, in estimating parameters and elasticities, therefore we generated *unrestricted* demand and income elasticities.

By manipulating the McDonald and Moffitt (1980) decomposition from equation (6) the expression in equation (20) can be used to show the change in probability of being above the limit (which is zero in this analysis) (in other words the probability of purchase), for the conditional sample for consumption of each beverage category due to a change in each explanatory variable, i.e. $\left(\frac{\partial F(z)}{\partial X}\right)$.

(20) $$\left(\frac{\partial F(z)}{\partial X}\right) = \frac{1}{E(y^*)} \left(\frac{\partial E(y)}{\partial X} - F(z)\frac{\partial E(y^*)}{\partial X}\right)$$

5. DATA

In order to complete this analysis, 2011 Nielsen Homescan panel data were used. This data set provides detailed purchase and demographic information of 62,092 households across the United States. This micro-level data is a representative sample of households across the country. The Nielsen Homescan data includes information collected from households that re-scan their purchases from retail outlets such as, grocery stores, department stores, convenience stores, drug stores, and club stores. This process allows the panelists to record the total expenditure and quantity information for all of the items they purchased for at-home use. The quantity and expenditure information is recorded alongside the demographic information and organized by household code.

The almond milk brands used in this study are WhiteWave Foods brand name Silk and Blue Diamond Growers brand name Almond Breeze. The soymilk brand used was WhiteWave Foods brand name Silk. The white milk category is comprised of conventional milk, with all organic and other specialty type milk products removed.

The demographic characteristics included in this study are household size, household income, age of household head, employment status of household head, education status of household head, region, race, Hispanic origin, age and presence of children, and the gender of the household head. Table 4 provides a detailed breakdown of the explanatory variables used in this study. In the data set, only expenditure data (in terms of dollars) and quantity data (in terms of ounces) are collected. Since no price is collected we must calculate a unit value and then use that as a proxy for price in the model. We calculated this unit value (prices in dollar per ounce) by taking the ratio of expenditure (dollars) to volume (ounces). The mean price for almond milk, soymilk, white milk, and lactose-free milk are $0.053/ounce, $0.055/ounce, $0.030/ounce, and $0.057/ounce, respectively.

Table 4 Summary Statistics of the Variables Used in the Model

Variable	Mean	Standard Deviation
Price of almond milk ($/ounce)	0.0530	0.0130
Price of soymilk ($/ounce)	0.0547	0.0167
Price of conventional white milk ($/ounce)	0.0300	0.0125
Price of lactose-free milk ($/ounce)	0.0565	0.0113
Household income (in ‘1000 dollars)	58.3154	31.9279
Age of Household Head less than 25 years (Base category)	0.0018	0.0424
Age of household head 25-29	0.0177	0.1319
Age of household head 30-34	0.0377	0.1905
Age of household head 35-44	0.1474	0.3545
Age of household head 45-54	0.2762	0.4471
Age of household head 55-64	0.2975	0.4572
Age of household head 65 or older	0.2218	0.4154
Household Head not employed for full pay (Base category)	0.4319	0.4953
Employment status part-time	0.1783	0.3828
Employment status full-time	0.3898	0.4877
Education of Household Head: Less than high school (Base category)	0.0239	0.1529
Education: High School	0.2374	0.4255
Education: Undergraduate	0.6183	0.4858
Education: Post-College	0.1204	0.3254
Race White (Base category)	0.8377	0.3688
Black	0.0938	0.2916
Asian	0.0286	0.1666
Other (non-Black, non-White, non-Asian)	0.0399	0.1958
Non-Hispanic Ethnicity (Base category)	0.9489	0.2201
Hispanic	0.0511	0.2201
Region: Pacific (Base category)	0.1230	0.3284

Table 4 Continued

Variable	Mean	Standard Deviation
New England	0.0454	0.2082
Middle Atlantic	0.1306	0.3370
East North Central	0.1811	0.3851
West North Central	0.0863	0.2809
South Atlantic	0.1982	0.3986
East South Central	0.0600	0.2375
West South Central	0.1023	0.3030
Mountain	0.0731	0.2603
No Child less than 18 years (Base category)	0.7869	0.4095
Age and Presence of Children less than 6-years	0.0275	0.1636
Age and Presence of Children between 6-12 years	0.0524	0.2227
Age and Presence of Children between 13-17 years	0.0668	0.2497
Age and Presence of Children less than 6 and 6-12	0.0244	0.1543
Age and Presence of Children less than 6 and 13-17	0.0041	0.0641
Age and Presence of Children between 6-12 and 13-17	0.0330	0.1786
Age and Presence of Children less than 6, 6-12 and 13-17 years	0.0049	0.0696
Household Head both Male and Female (Base category)	0.6540	0.4757
Female head only	0.2500	0.4331
Male head only	0.0959	0.2945

Source: Nielsen Homescan data 2011, calculated by the author, base category of dummy variables is printed in *italics*

Household income takes on a wide range of values, from households that make less than $5,000 per year to households that earn over $100,000 per year. The mean household income for this data set was $58,315 (Table 4), most households fall into the $50,000 to $59,999 income range for the 2011 calendar year.

Table 4 shows the age of household head categories, as well as the base category used in the model. Households whose household head was under the age of 25 were used as the base category. Households with a household head between 25 and 29 made up

1.8% of the sample. Households with a household head between 30 and 34 made up another small portion of the sample with 3.8% of households falling in that category. Households with a household head between 35 and 44 represented 14.8% of the sample. Households with a household head between 45 and 54 were the second largest group represented, with 27.6% of the households falling into that category. The largest age group represented was households with a household head between 55 and 64 which made up 29.7% of the sample. Households with a household head over 65 years old accounted for 22.1% of the sample.

Employment status was broken into three categories depending on how many hours the male or female household head worked per week. Households with a household head that were not employed for full pay were used as the base category. Households with a household head who worked under 34 hours per week were given a part time employment status and made up 17.8% of the sample. Households where the household head worked at least 35 hours per week were given a full time employment status and accounted for 39% of the sample.

The level of education of each household head also was considered. Those households where the household head did not graduate from high school were used as the base category. Households where the household head graduated high school but did not attend college made up 23.7% of the sample. Households with a household head that attended some college or graduated college made up 61.8% of the sample. Households with a household head that earned post-college education made up 12% of the sample.

The categories of White, Black, Asian, and Other were used to classify race. Those household heads who classified as white were used as the base category. Household heads who classified as Black made up 9.4% of the sample. Asian household heads accounted for 2.9% of the sample. Household heads who classified as other made up 4% of the sample. Hispanic-origin was also considered. Households whose household head is Hispanic accounted for 5.1% of the sample.

Traditionally, region in which the household is located is broken out into four basic categories, East, Midwest, West, and South. In this study, region was broken out into nine distinct categories. The breakdown is displayed below in Table 5.

Table 5 United States Census Bureau Regions and States

<table>
<tr><th colspan="6">EAST</th></tr>
<tr><td colspan="3"><u>New England</u>

Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont</td><td colspan="3"><u>Middle Atlantic</u>

New Jersey, New York, Pennsylvania</td></tr>
<tr><th colspan="6">CENTRAL</th></tr>
<tr><td colspan="3"><u>East North Central</u>

Indiana, Illinois, Michigan, Ohio, Wisconsin</td><td colspan="3"><u>West North Central</u>

Iowa, Nebraska, Kansas, North Dakota, South Dakota, Minnesota, Missouri</td></tr>
<tr><th colspan="6">SOUTH</th></tr>
<tr><td colspan="2"><u>South Atlantic</u>

Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia</td><td colspan="2"><u>East South Central</u>

Alabama, Kentucky, Mississippi, Tennessee</td><td colspan="2"><u>West South Central</u>

Arkansas, Louisiana, Oklahoma, Texas</td></tr>
<tr><th colspan="6">WEST</th></tr>
<tr><td colspan="3"><u>Mountain</u>

Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming</td><td colspan="3"><u>Pacific</u>

Alaska, California, Hawaii, Oregon, Washington</td></tr>
</table>

Source: U.S. Department of Commerce Economics and Statistics Administration U.S. Census Bureau

The Pacific region is the base category. Household heads from New England make up 4.5% of the sample, 13.1% are from the Middle Atlantic, 18.1% are from the East North Central, 8.6% come from the West North Central, 19.8% are from the South Atlantic, 6% are from the East South Central, and 10.2% are from West South Central, and 7.3% are from the states in the Mountain region. The distribution of households in this data set is consistent with the U.S. Census Bureau calculated distribution for each region.

The age and presence of children variable gives insight into not only the presence of children but their respective ages as well. This variable is broken down into eight categories based on the age of the children. Households with no children were used as the base category. Households with children less than six years old made up 2.8% of the sample. Households with children between the ages of 6 and 12 made up 5.2%. Households with children between the ages of 13 and 17 made up 6.7%. Households that had children under the age of six and children between the ages of 6 and 12 made up 2.4% of the sample. Households with children under the age of six and children between the ages of 13 and 17 made up less than 1% of the sample. Households with children between the ages of 6 and twelve, as well as children between the ages of 13 and 17 made up 3.3% of the sample. Finally, households with children in all age categories made up less than 1% of the sample.

Gender of the household head was classified into three categories. The base category included households with both a male and female household head. Households with a female household head only made up 25% of the sample, and households with a male household head only made up 9.6% of the sample.

6. EMPIRICAL RESULTS

Summary statistics for market penetration (ratio of number of households that purchased the beverage to the total number of households sampled), price (unit value), expenditure, and quantity are displayed in Table 6.

Table 6 Summary Statistics of Market Penetration, Price, Expenditure and Quantity for Almond Milk, Soymilk, Conventional White Milk, and Lactose-free Milk

	Almond Milk	Soymilk	Conventional White Milk	Lactose-free Milk
Market Penetration	12.06%	10.91%	92.72%	7.24%
Unconditional Average Price ($/ounce)	0.0531	0.0548	0.0301	0.0560
Conditional Average Price ($/ounce)	0.0530	0.0547	0.0300	0.0566
Unconditional Average Expenditure ($/HH/Year)	2.59	2.59	77.40	3.21
Conditional Average Expenditure ($/HH/Year)	21.51	23.74	83.47	44.29
Average Unconditional Quantity (oz/HH/Year)	51.19	52.39	2991.90	57.89
Average Conditional Quantity (oz/HH/Year)	424.56	480.04	3226.68	799.79

Source: Calculated by the author

Out of the 62,092 households sampled, 7,487 purchased almond milk, 6,776 purchased soymilk, 57,574 purchased conventional white milk, and 4,494 purchased lactose-free milk. This results in a market penetration of 12.06%, 10.91%, 92.72%, and 7.24% for almond milk, soymilk, conventional white milk, and lactose free milk, respectively. The unconditional average prices for almond milk, soymilk, conventional white milk, and lactose-free milk are respectively $0.053/ounce, $0.054/ounce,

$0.030/ounce, and $0.056/ounce. The average price paid by households that purchased almond milk was $0.05 per ounce ($3.39 for 64 ounces (half a gallon); the most popular container size). The average price paid by households that purchased soymilk was $0.05 per ounce ($3.50 for 64 ounces). The average price paid by households that purchased conventional white milk was $0.03 per ounce ($1.92 for 64 ounces). The average price paid by households that purchased lactose-free milk was $0.06 per ounce ($3.62 for 64 ounces).

The average expenditure (expressed in terms of dollars spent per household for the calendar year 2011) for all households considered in this study for almond milk, soymilk, conventional white milk, and lactose-free milk respectively was $2.59, $2.59, $77.40 and $3.21. The conditional average expenditure for almond milk, soymilk, conventional white milk, and lactose-free milk respectively was $21.51, $23.74, $83.47, and $44.29. As to be expected, the conditional average expenditure is much larger than the unconditional average expenditure because they are households that actually purchased each beverage considered in 2011.

The average consumption/purchase of almond milk by a consuming household (conditional sample) was calculated to be 431 ounces per year (approximately seven, half-gallon containers per household per year). The average consumption/purchase of soymilk by a consuming household was calculated to be 480 ounces per year (approximately eight, half-gallon containers per household per year). The average consumption/purchase of conventional white milk by a consuming household was calculated to be 3,227 ounces (approximately fifty half-gallon containers per household per year). The average consumption/purchase of lactose-free milk by a consuming household was calculated to be 800 ounces per year (approximately twelve, half-gallon containers per household per year).

The average almond milk consumption/purchase of all households (unconditional sample) in this study was calculated to be 51 ounces (less than one half gallon container per household per year). The average soymilk consumption/purchase of all households in this study was calculated to be 52 ounces (less than one half gallon container per household per year). The average conventional white milk consumption was calculated to be 2,991 ounces (approximately 47, half-gallon containers per household per year). The average lactose-free milk consumption/purchase was calculated to be 58 ounces (less than one half-gallon container per household per year).

The *chi-square* test results are displayed in Table 7. These results show the joint significance for the demographic variables considered in each single-equation Tobit model for the four beverage categories. The level of significance used in conducting these tests is 0.05. In the almond milk equation, age of household head, employment status, level of education, race, region, age and presence of children, and the gender of the household head, are significant. In the soymilk equation, the age a presence of children is not significant as a demographic variable. In the conventional white milk equation, all of the demographic variables are significant when taken as a group. The results for the lactose-free milk equation also show that each group of demographic variables is significant when taken as a whole.

Table 8 displays the single-equation Tobit regression results for almond milk, soymilk, conventional white milk, and lactose-free milk. The economic determinants and demographic variables are considered significant if they have a *p*-value equal to or less than 0.05.

The significant economic determinants of demand for almond milk are the price of soymilk, the price of almond milk, the price of conventional white milk, the price of lactose-free milk, and household income. Taken individually, the age of the household head did not have a significant effect on the demand for almond milk. The significant demographic variables impacting demand for almond milk include, employment status, education, race, Hispanic origin, region, and gender. The presence of a teenager (age 13-17) and a pre-teen (age 6-12) in the household also has a significant impact on the demand for almond milk, whereas the other age categories for children were not significant.

The significant determinants of demand for soymilk are the price of soymilk, price of almond milk, price of conventional white milk, price of lactose-free milk and household income. The significant demographic variables include, age of household head, employment status, education, race, Hispanic origin, region, and gender of the household head. The age and presence of children is not a significant determinant of demand for soymilk, a result that is complemented by the *chi-square* test results.

The significant economic determinants of demand for conventional white milk are the price of almond milk, price of white milk, and household income. The significant demographic variables include employment status, education, race, Hispanic origin, region, age and presence of children, and gender.

The significant economic determinants of demand for lactose-free milk are price of almond milk, price of conventional white milk, price of lactose-free milk, and household income. Employment status, education, race, Hispanic origin, and region are the significant demographic variables impacting the demand for lactose-free milk.

Table 9 presents the conditional marginal effects. Table 10 shows the probability of being above the limit (or probability of purchase) for each demographic variable for each beverage category. Median values were used to reduce the impact of outliers and the possibility of skewed data.

Table 7 Chi-square Test of Joint Significance of Demographic Variables Considered in the Almond Milk, Soymilk, Conventional White Milk, and Lactose-free Milk Tobit Models

	Label	Almond Milk	Soymilk	Conventional White Milk	Lactose-free Milk
Age of Household Head	agehh2529 = 0 agehh3034 = 0 agehh3544 = 0 agehh4554 = 0 agehh5564 = 0 agehhgt64 = 0	<.0001	0.0002	<.0001	0.0009
Employment Status	emphhpt=0 emphhft=0	<.0001	0.0003	<.0001	<.0001
Education	eduhhhs = 0 eduhhu = 0 eduhhpc = 0	<.0001	<.0001	<.0001	<.0001
Race	black = 0 asian = 0 other = 0	<.0001	<.0001	<.0001	<.0001
Region	newengland = 0 middleatlantic = 0 eastnorthcentral = 0 westnorthcentral = 0 southatlantic = 0 eastsouthcentral = 0 westsouthcentral = 0 mountain = 0	<.0001	<.0001	<.0001	<.0001
Age and Presence of Children	aclt6_only = 0 ac6_12only = 0 ac13_17only = 0 aclt6_6_12only = 0 aclt6_13_17only = 0 ac6_12and13_17only = 0 aclt6_6_12and13_17 = 0	0.0233	0.1410	<.0001	0.0039
Gender	fhonly = 0 mhonly = 0	<.0001	<.0001	<.0001	<.0001

Source: Calculated by the author

Table 8 Tobit Regression Results for Almond Milk, Soymilk, White Milk, and Lactose-free Milk

	Almond Milk			**Soymilk**			**Conventional White Milk**			**Lactose-free Milk**		
Variable	Estimate	Std Error	p-Value	Estimate	Std Error	p-Value	Estimate	Std Error	p-Value	Estimate	Std Error	p-Value
Intercept	**-8361.9**	472.93	<.0001	**-9445.2**	586.06	<.0001	**-11160**	216.36	<.0001	**-9839.2**	993.78	<.0001
Log price of soymilk	**-536.44**	76.78	<.0001	**-1737.6**	68.27	<.0001	121.53	158.61	0.4436	-213.43	187.68	0.2555
Log price of almond milk	**-1237.5**	65.11	<.0001	**-1058.2**	109.19	<.0001	**-546.03**	175.92	0.0019	**-999.3**	220.46	<.0001
Log price of white milk	**459.34**	25.42	<.0001	**506.91**	31.40	<.0001	**-3712.9**	47.26	<.0001	**1190.7**	57.85	<.0001
Log price of lactose-free milk	**-1046.3**	122.08	<.0001	**-549.81**	156.54	0.0004	301.92	208.36	0.1473	**-2391.2**	176.78	<.0001
Log household income	**117.22**	13.36	<.0001	**86.02**	16.33	<.0001	**50.85**	23.41	0.0299	**190.20**	29.69	<.0001
Age of household head 25-29	179.42	179.69	0.318	**-375.58**	184.58	0.0419	232.22	293.04	0.4281	-419.83	384.74	0.2752
Age of household head 30-34	103.49	176.54	0.5577	**-454.90**	179.71	0.0114	512.08	284.98	0.0724	-378.01	374.27	0.3125
Age of household head 35-44	90.17	174.20	0.6047	**-477.11**	176.03	0.0067	522.87	279.04	0.061	-354.07	367.73	0.3356
Age of household head 45-54	-23.85	173.78	0.8909	**-515.74**	175.31	0.0033	**741.78**	278.34	0.0077	-333.07	366.56	0.3635
Age of household head 55-64	-72.36	173.74	0.677	**-534.49**	175.27	0.0023	**581.23**	277.86	0.0365	-227.83	366.42	0.5341
Age of household head 65 or older	-189.92	174.16	0.2755	**-595.10**	175.89	0.0007	506.56	278.20	0.0686	-108.00	367.18	0.7687
Employment status part-time	**66.60**	20.78	0.0014	**68.35**	25.65	0.0077	**-167.49**	38.74	<.0001	-91.36	48.10	0.0575
Employment status full-time	**-65.98**	18.83	0.0005	-36.04	23.15	0.1195	**-347.93**	34.50	<.0001	**-215.14**	42.96	<.0001
Education high school	101.1	56.70	0.0746	-5.02	64.84	0.9383	-107.58	90.06	0.2323	-6.96	118.74	0.9532
Education undergraduate	**256.29**	55.53	<.0001	**139.99**	63.35	0.0271	**-240.40**	88.41	0.0065	**253.28**	115.86	0.0288
Education post-college	**298.28**	58.97	<.0001	**200.97**	67.87	0.0031	**-305.27**	96.39	0.0015	**383.36**	123.94	0.002
Black	**127.22**	23.99	<.0001	**264.91**	28.95	<.0001	**-1363.9**	47.45	<.0001	**752.05**	50.17	<.0001
Asian	**172.88**	39.61	<.0001	**365.950**	47.27	<.0001	**-825.54**	82.44	<.0001	**488.23**	89.37	<.0001
Other	73.56	37.99	0.0529	**152.53**	45.74	0.0009	**-486.30**	74.22	<.0001	**373.42**	82.076	<.0001
Hispanic	**112.71**	33.43	0.0007	**201.95**	40.25	<.0001	**-152.41**	66.20	0.0213	**479.79**	72.16	<.0001

Table 8 Continued

	Almond Milk			Soymilk			Conventional White Milk			Lactose-free Milk		
Variable	Estimate	Std Error	p-Value	Estimate	Std Error	p-Value	Estimate	Std Error	p-Value	Estimate	Std Error	p-Value
New England	**-232.73**	42.13	<.0001	**-161.73**	52.55	0.0021	**415.43**	77.48	<.0001	-6.60	86.40	0.9391
Middle Atlantic	**-100.63**	28.88	0.0005	**-74.61**	35.80	0.0371	**216.64**	54.46	<.0001	76.29	63.15	0.227
East North Central	**-312.02**	32.40	<.0001	**-200.90**	40.12	<.0001	**240.30**	57.54	<.0001	**-618.19**	69.04	<.0001
West North Central	**-329.54**	36.19	<.0001	**-269.66**	44.55	<.0001	**710.39**	63.67	<.0001	**-790.06**	84.58	<.0001
South Atlantic	**-210.33**	28.82	<.0001	**-231.10**	36.31	<.0001	**387.12**	53.57	<.0001	**-224.26**	61.99	0.0003
East South Central	**-382.2**	40.37	<.0001	**-334.74**	50.29	<.0001	**342.40**	71.81	<.0001	**-850.17**	93.38	<.0001
West South Central	**-480.06**	40.49	<.0001	**-305.68**	50.25	<.0001	-36.62	68.78	0.5944	**-687.94**	80.71	<.0001
Mountain	-62.79	34.86	0.0717	-67.86	43.72	0.1206	**-175.49**	64.93	0.0069	**-442.68**	83.81	<.0001
Children less than 6 years	-39.84	45.32	0.3793	84.35	54.80	0.1237	**1520.41**	87.57	<.0001	**264.20**	102.61	0.01
Children 6-12 years	-34.93	33.79	0.3013	61.35	40.99	0.1345	**1116.89**	63.87	<.0001	27.39	79.72	0.7312
Children13-17 years	**-60.58**	30.49	0.0469	28.90	36.92	0.4338	**1500.01**	56.06	<.0001	-107.27	73.15	0.1425
Children under 6 and 6-12 years	-16.92	47.16	0.7197	-78.26	60.41	0.1951	**1930.75**	92.59	<.0001	**230.93**	109.89	0.0356
Children under 6 and 13-17	-153.95	112.45	0.171	-23.80	136.75	0.8618	**1951.94**	207.56	<.0001	-587.74	308.67	0.0569
Children 6-12 and 13-17 years	**-151.32**	43.17	0.0005	-79.75	52.94	0.1319	**2368.03**	78.58	<.0001	-166.3	105.26	0.1142
Children under 6, 6-12, 13-17	-82.17	99.15	0.4073	-8.09	122.30	0.9473	**2967.88**	192.77	<.0001	-55.43	249.37	0.8241
Female head only	**52.38**	19.64	0.0076	-45.07	24.38	0.0645	**-1145.04**	35.87	<.0001	-71.67	44.32	0.1059
Male head only	**-220.67**	29.29	<.0001	**-204.69**	35.11	<.0001	**-1034.74**	49.06	<.0001	**-299.19**	63.96	<.0001
Sigma	**1115.2**	10.4	<.0001	**1331.4**	12.98	<.0001	**3277.09**	9.75	<.0001	**2149.9**	26.53	<.0001
Model Fit Statistics												
R-squared	0.0149			0.0134			0.2181			0.0096		

Source: Calculated by the author
Note: Std Error is abbreviation for Standard Error

Table 9 Median Conditional Marginal Effects of the Respective Explanatory Variables in Soymilk, Almond Milk, Lactose-Free Milk and Conventional White Milk Demand Equation

Variable	Almond Milk	Soymilk	Conventional White Milk	Lactose-free Milk
Age of household head 25-29	33.51	-68.45	141.04	-67.80
Age of household head 30-34	19.32	-82.90	311.02	-61.05
Age of household head 35-44	16.84	-86.95	317.57	-57.18
Age of household head 45-54	-4.45	-93.99	450.53	-53.79
Age of household head 55-64	-13.51	-97.41	353.02	-36.79
Age of household head >65	-35.47	-108.45	307.67	-17.44
Employment status part-time	12.44	12.46	-101.73	-14.75
Employment status full-time	-12.32	-6.57	-211.32	-34.74
Education: high school	18.88	-0.91	-65.34	-1.12
Education: undergraduate	47.86	25.51	-146.01	40.90
Education: post-college	55.70	36.62	-185.41	61.91
Black	23.76	48.28	-828.41	121.45
Asian	32.28	66.69	-501.41	78.84
Other	13.74	27.80	-295.36	60.30
Hispanic	21.05	36.80	-92.57	77.48
New England	-43.46	-29.47	252.32	-1.07
Middle Atlantic	-18.79	-13.60	131.58	12.32
East North Central	-58.27	-36.61	145.95	-99.83
West North Central	-61.54	-49.14	431.47	-127.59
South Atlantic	-39.28	-42.12	235.12	-36.22
East South Central	-71.38	-61.00	207.96	-137.29
West South Central	-89.65	-55.71	-22.24	-111.53
Mountain	-11.73	-12.37	-106.59	-71.49
Children less than 6 years	-7.44	15.37	923.45	42.67
Children 6-12 years	-6.52	11.18	678.36	4.42
Children 13-17 years	-11.31	5.27	911.05	-17.32
Children < 6 & 6-12 years	-3.16	-14.26	1172.67	37.29
Children <6 & 13-17 years	-28.75	-4.34	1185.54	-94.91
Children 6-12 & 13-17 years	-28.26	-14.53	1438.26	-26.86
Children <6 & 6-12 & 13-17	-15.34	-1.47	1802.59	-8.95
Female head only	9.78	-8.21	-695.46	-11.57
Male head only	-41.21	-37.30	-628.47	-48.32

Source: Calculated by the author

Note: Conditional marginal effects are in liquid ounces; 16 liquid ounces equal to one gallon.

Table 10 Median Change in the Probability of Purchasing Almond Milk, Soymilk, Conventional White Milk and Lactose-free Milk for the Respective Explanatory Variables

Variable	Almond Milk	Soymilk	Conventional White Milk	Lactose-free Milk
Age of household head 25-29	0.0347	-0.0580	0.0184	-0.0308
Age of household head 30-34	0.0200	-0.0702	0.0406	-0.0277
Age of household head 35-44	0.0174	-0.0737	0.0415	-0.0260
Age of household head 45-54	-0.0046	-0.0796	0.0588	-0.0244
Age of household head 55-64	-0.0140	-0.0825	0.0461	-0.0167
Age of household head >65	-0.0367	-0.0919	0.0402	-0.0079
Employment status part-time	0.0129	0.0106	-0.0133	-0.0067
Employment status full-time	-0.0128	-0.0056	-0.0276	-0.0158
Education: high school	0.0195	-0.0008	-0.0085	-0.0005
Education: undergraduate	0.0495	0.0216	-0.0191	0.0186
Education: post-college	0.0577	0.0310	-0.0242	0.0281
Black	0.0246	0.0409	-0.1081	0.0551
Asian	0.0334	0.0565	-0.0654	0.0358
Other	0.0142	0.0235	-0.0386	0.0274
Hispanic	0.0218	0.0312	-0.0121	0.0352
New England	-0.0450	-0.0250	0.0329	-0.0005
Middle Atlantic	-0.0195	-0.0115	0.0172	0.0056
East North Central	-0.0603	-0.0310	0.0191	-0.0453
West North Central	-0.0637	-0.0416	0.0563	-0.0579
South Atlantic	-0.0407	-0.0357	0.0307	-0.0164
East South Central	-0.0739	-0.0517	0.0271	-0.0623
West South Central	-0.0928	-0.0472	-0.0029	-0.0504
Mountain	-0.0121	-0.0105	-0.0139	-0.0324
Children less than 6 years	-0.0077	0.0130	0.1205	0.0194
Children 6-12 years	-0.0068	0.0095	0.0885	0.0020
Children 13-17 years	-0.0117	0.0045	0.1189	-0.0079
Children < 6 & 6-12 years	-0.0033	-0.0121	0.1531	0.0169
Children <6 & 13-17 years	-0.0298	-0.0037	0.1547	-0.0431
Children 6-12 & 13-17 years	-0.0293	-0.0123	0.1877	-0.0122
Children <6 & 6-12 & 13-17	-0.0159	-0.0012	0.2353	-0.0041
Female head only	0.0101	-0.0066	-0.0908	-0.0053
Male head only	-0.0427	-0.0316	-0.0820	-0.0219

Source: Calculated by the author

As shown in Table 9, for almond milk, households where the household head is employed part-time consume 12.44 ounces more than households where the household head is not employed. Whereas, households where the household head is employed full-time consume 12.32 ounces less than households where the household head is not employed. The average change in probability of consumption (Table 10) for a change in employment status at the median is 0.0129 and -0.0128 for part-time and full-time respectively. This means that households where the household head is employed part-time is 1.29% more likely to consume almond milk than the base category of households where the household head is not employed for full pay. It then follows that households where the household head is employed full-time are 1.28% less likely to consume almond milk than households where the household head is not employed for full pay.

The higher level of education of the household head, the more likely households are to consume almond milk. Households where the household head has a college education consume 47.86 more ounces and are approximately 5% more likely to consume almond milk than households with less than a high school education. Post-college educated households consume 55.70 more ounces and are 5.8% more likely to consume almond milk than the base category.

Compared to the base category of white households, households where the household head identifies as black are 2.5% more likely to consume almond milk and consume 23.76 more ounces. Households where the household head identifies as Asian are 3.3% more likely to consume almond milk and consume 32.28 more ounces than white households. Households where the household head has a Hispanic origin are 2.2% more likely to consume almond milk and consume 21.05 ounces more than non-Hispanic households.

Region was broken down into nine categories with Pacific as the base. All, except Mountain (which was marginally significant) were significant determinants of demand for almond milk. Households in the New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, and West South Central regions were less like to consume almond milk than households in the Pacific region. These households consume between 18.79 to 89.65 less ounces than the base category, with a range of probabilities from 1.9% to 9.3%.

Households with children ages 13 to 17 years old are 1.2% less likely to consume almond milk compared to households with no children. These households consume 11.31 less ounces than those without children. Households that have children that are between the ages of 6 and 12, as well as children between the ages of 13 and 17 are approximately 3% less likely to consume almond milk and consume 28.26 less ounces than households without children.

For soymilk, the older the household head is, the less likely they are to consume the beverage. Households with a household head aged 25 to 29 consume 68.45 less ounces than households where the household head is younger than 25 years old. As the age of the household head increases, the less likely they are to consume soymilk with households where the household head is older than 65 years old consuming over 100 ounces less than the base category. Households with a household head aged 25 to 29 are approximately 6% less likely to consume soymilk than the base category and that percentage only increases as the age increases, maxing out at 9.2% for households where the household head is older than 65 years.

Households where the household head is employed part-time are approximately 1% more likely to consume soymilk than the base category of households who are not employed for full pay. These households consume 12.46 more ounces of soymilk than the base category.

College and post-college educated households are more likely to consume soymilk than the base category of households who have less than a high school education. College educated households are 2.2% more likely to consume soymilk and consume 25.51 more ounces than the base category. Post-college educated households are 3.1% more likely to consume soymilk and consume 36.62 more ounces than the base category.

All race categories identified consume more soymilk than the white base category. Households that identified as black consume 48.28 more ounces than white households, and are 4.1% more likely to consume soymilk. Households that identified as Asian are 5.7% more likely to consume soymilk and consume 66.69 more ounces than households that identified as white. Households that didn't identify as black, white, or Asian were classified as other. These households consume 27.80 more ounces and are 2.4% more likely to consume soymilk than the base category. Households where the household head is from Hispanic origin were 3.1% more likely to consume soymilk than non-Hispanic households.

As with almond milk, households in the base category of the Pacific were the most likely to consume soymilk. Again, the Mountain region was not significant. Households in the New England, Middle Atlantic, East North Central, West North

Central, South Atlantic, East South Central, and West South Central regions were anywhere from 1.1% to 5.2% less likely to consume than households in the Pacific. Households located in the East South Central region are the least likely to consume soymilk and consume 61.00 ounces less than the base category. The other regions consume anywhere between 13.60 to 55.71 ounces less than households in the Pacific region.

The age and presence of children was not a significant determinant of demand for soymilk. Although, households with a male household head are 3.2% less likely to consume soymilk and consume 37.30 ounces less than the base category of households with both a male and female household head.

For white milk, households where the household head was between the ages of 45 and 54 consume 450.53 more ounces than households where the household head was less than 25. In fact, these households were approximately 6% more likely to consume white milk than the base category. Households where the household head was between 55 and 64 also consume more white milk than younger households. They consume 353.02 more ounces and were 4.6% more likely to consume conventional white milk.

Households where the household head was not employed for full pay consume the most white milk. Compared to households where the household head was employed part-time, who were 1.3% less likely to consume soymilk. These households where the head was employed part-time consume 101.73 ounces less than the base category of households where the head is not employed for full pay. If the household head is employed full-time, they are even less likely to consume white milk. The households where the household head is employed full-time are 2.8% less likely to consume conventional white milk and they consume 211.32 ounces less than the base category.

Unlike the dairy alternative beverages previously discussed, the more education a household has, the less likely they are to consume conventional white milk. College educated households consume 146.01 less ounces of conventional white milk than households with less than a high school education. Post-college educated households are 2.4% less likely to consume conventional white milk and consume 185.41 less ounces than households in the base category.

When compared to dairy alternative beverages, race has the opposite effect on conventional white milk. Households that identified as anything other than white are less likely to consume white milk. Households that identified as black are 10.8% less likely to consume conventional white milk than the white base category. Black households consume 828.41 ounces less than the base category. Households that identified as Asian are also less likely to consume conventional white milk. They consume 501.41 ounces less than white households and are 6.5% less likely to consume conventional white milk. Households that identified as other are approximately 3.9% less likely to consume conventional white milk than the base category. Households where the household head is of Hispanic origin are 1.2% less likely to consume conventional white milk than non-Hispanic households. Hispanic households consume 92.57 less ounces of conventional white milk than non-Hispanic households.

Unlike the dairy alternative beverages, households in the New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, and Mountain regions are more likely to consume conventional white milk than households in the Pacific. Households located in the West North Central region consume the most, consuming 431.47 more ounces than households in the Pacific region. Households in these regions are anywhere from 1.7% to 5.6% more likely to consume white milk than the base category. Conversely, households in the Mountain region are 1.4% less likely to consume conventional white milk and consume 106.59 ounces less than households in the Pacific region.

Households with children are more likely to consume conventional white milk than households with no children. Households with children in all three of the age categories, less than 6 years old, between 6 and 12, and between 13 and 17, consume 1802.59 more ounces than households without children. These households are 23.5% more likely to consume conventional white milk than the base category. Households with children under 6 years of age are 12.1% more likely to consume white milk than the base category and consume 923.45 more ounces. Households with children between 6 and 12 are the least likely to consume conventional white milk, but they are still 8.9% more likely to consume white milk than households without children. Overall, households with children are anywhere between 8.9% and 23.5% more likely to consume conventional white milk than households without children. They consume anywhere between 678.36 ounces and 1802.59 more ounces than the base category.

Households with a female head only are 695.46 ounces less than households with both a male and female household head. These household are 9.1% less likely to consume conventional white milk than the base category. Households with a male household head only are also less likely to consume conventional white milk, and consume 629.47 ounces less than the base category.

For lactose-free milk, age was not a significant determinant of demand. Households with household heads who are employed full-time are 1.6% less likely to consume lactose-free milk. These households consume 34.74 ounces less than the base category of households with household heads who are not employed for full pay. College and post-college educated households consume 40.90 and 61.91 ounces more lactose-free milk, respectively, than households with less than a high school education. College educated households are 1.9% more likely to consume lactose-free milk than the base category, and post-college educated households are 2.8% more likely to consume lactose-free milk than the base category.

Lactose-free milk is more likely to be consumed in non-white households. Households that identified as black are 5.5% more likely to consume lactose-free milk than households that identified as white. These households consume 121.45 more ounces of lactose-free milk than the base category. Households that identified as Asian are 3.6% more likely to consume lactose-free milk and consume 78.84 more ounces than households with white household heads. Households that identified as other consume 60.30 more ounces of lactose-free milk than white households. These households are 2.7% more likely to consume lactose-free milk than the base category. Hispanic households are also more likely to consume lactose-free milk, being 3.5% more likely than non-Hispanic households. Households where the household head is Hispanic consume 77.48 more ounces than non-Hispanic households.

Households in the East North Central, West North Central, South Atlantic, East South Central, West South Central, and Mountain regions are less likely to consume lactose-free milk. Households in this region consume between 36.22 and 137.29 ounces less than households in the Pacific region. These households are anywhere between 1.6% and 6.2% less likely to consume lactose-free milk than households in the base category.

Households with children under 6 years of age are 1.9% more likely to consume lactose-free milk. These households with young children consume 42.67 more ounces of lactose-free milk than households with no children. Households with children under 6 and children between 6 and 12 years of age are 1.7% more likely to consume lactose-free milk than the base category. These households consume 37.29 more ounces than households without children.

Households with a male household head only are 2.2% less likely to consume lactose-free milk than households with both a male and female household head. These households consume 48.32 ounces less than the base category.

Table 11 reports the median unconditional marginal effects for each demographic variable. As to be expected, the unconditional marginal effects are smaller than the conditional marginal effects in terms of absolute value.

Table 11 Median Unconditional Marginal Effects of the Respective Explanatory Variables in the Almond Milk, Soymilk, Conventional White Milk and Lactose-free Milk Demand Equation

Variable	Almond Milk	Soymilk	Conventional White Milk	Lactose-free Milk
Age of household head 25-29	18.79	-34.76	188.92	-24.72
Age of household head 30-34	10.84	-42.10	416.60	-22.25
Age of household head 35-44	9.44	-44.16	425.38	-20.84
Age of household head 45-54	-2.50	-47.73	603.47	-19.61
Age of household head 55-64	-7.58	-49.46	472.86	-13.41
Age of household head >65	-19.89	-55.07	412.11	-6.36
Employment status part-time	6.97	6.33	-136.26	-5.38
Employment status full-time	-6.91	-3.34	-283.05	-12.67
Education: high school	10.59	-0.46	-87.52	-0.41
Education: undergraduate	26.84	12.96	-195.58	14.91
Education: post-college	31.24	18.60	-248.35	22.57
Black	13.32	24.52	-1109.63	44.27
Asian	18.10	33.87	-671.62	28.74
Other	7.70	14.12	-395.63	21.98
Hispanic	11.80	18.69	-123.99	28.25
New England	-24.37	-14.97	337.97	-0.39
Middle Atlantic	-10.54	-6.91	176.24	4.49
East North Central	-32.68	-18.59	195.50	-36.39
West North Central	-34.51	-24.96	577.94	-46.51
South Atlantic	-22.03	-21.39	314.94	-13.20
East South Central	-40.02	-30.98	278.56	-50.05
West South Central	-50.27	-28.29	-29.79	-40.50
Mountain	-6.58	-6.28	-142.77	-26.06
Children less than 6 years	-4.17	7.81	1236.93	15.55
Children 6-12 years	-3.66	5.68	908.64	1.61
Children 13-17 years	-6.34	2.67	1220.33	-6.32
Children < 6 & 6-12 years	-1.77	-7.24	1570.75	13.60
Children <6 & 13-17years	-16.12	-2.20	1588.00	-34.60
Children 6-12 & 13-17 years	-15.85	-7.38	1926.51	-9.79
Children <6 & 6-12 & 13-17	-8.60	-0.75	2414.52	-3.26
Female head only	5.49	-4.17	-931.55	-4.22
Male head only	-23.11	-18.94	-841.81	-17.61

Source: Calculated by authors

Note: Unconditional marginal effects are in liquid ounces; 16 liquid ounces equal to one gallon

For almond milk, the unconditional marginal effect for the employment status variable for household heads who are employed part-time tell us that, these households would consume 6.97 more ounces than households where the household head is not employed for full pay. Households where the household head is employed full-time would consume 6.91 ounces less than the base category.

Household heads with an undergraduate degree would consume 26.84 more ounces of almond milk than households with less than a high school education. Households with a post-college educated household head would consume even more, consuming 31.24 ounces more almond milk than the base category.

More almond milk would be consumed by non-white households than white households. Households that identified as black would consume 13.32 more ounces than the base category of households that identified as white. Households that identified as Asian would also consume more almond milk than white households, consuming 18.10 more ounces. Households with a household head who was of Hispanic-origin would consume 11.80 more ounces than households where the household head was not of Hispanic-origin.

Households in the New England region would consume 24.37 ounces less than households in the Pacific region. Households in the Middle Atlantic region would consume 10.54 ounces less, the East North Central region would consume 32.68 ounces less, the West North Central region would consume 34.51 ounces less, the South Atlantic region would consume 22.03 ounces less, the East South Central region would consume 40.02 ounces less, and the West South Central region would consume 50.27 ounces less than the Pacific region.

Households with teenagers (children between the ages of 13 and 17 years old) would consume 6.34 ounces less than households with no children. Households with children between the ages of 6 and 12, as well as with children between the ages of 13 and 17 years old would consume 15.85 ounces less than the base category.

Households with a female household head only would consume 5.49 ounces more of almond milk than households with both a male and female household head. Households with a male household head only would consume 23.11 ounces less than the base category.

The soymilk results show that the older the household head is the less soymilk they would consume. Households where the household head was between the ages of 25 and 29 would consume 34.76 less ounces than households where the household head was younger than 25. Households where the household head is 30 to 34 years old would consume 42.10 ounces less, households where the household head is 35 to 44 years old would consume 44.16 ounces less, households where the household head is 45 to 54 years old would consume 47.73 ounces less, households where the household head is 55 to 64 years old would consume 49.46 ounces less, and households where the household head is older than 65 would consume 55.07 ounces less than the base category.

Households where the household head is employed part-time would consume 6.33 ounces more of soymilk than households where the household head is not employed for full pay. College and post-college educated households would consume more than households with less than a high school education, consuming 12.96 ounces and 18.60 ounces more respectively.

Households that identified as black would consume 24.52 ounces more than households that identified as white. Households that identified as Asian would consume more than any other race, consuming 33.87 ounces more than white households. Households that identified as other would consume 14.12 ounces more than the base category. Hispanic households would consume 18.69 ounces more than non-Hispanic households.

Households in the Pacific would consume more soymilk than households in any other region. Households in New England would consume 14.97 ounces less, households in the Middle Atlantic region would consume 6.91 ounces less, households in the East North Central region would consume 18.59 ounces less, households in the West North Central region would consume 24.96 ounces less, households in the South Atlantic region would consume 21.39 ounces less, households in the East South Central region would consume 30.98 ounces less, and households in the West South Central region would consume 28.29 ounces less than households in the base category.

Households with a male household head only would consume 18.94 ounces less than households in the base category that have both a male and female household head.

Conventional white milk would be predominately consumed in households with a household head that is between the ages of 45 and 54. These households would consume 603.47 more ounces than households with a household head who is younger than 25 years old. Households with a household head between the ages of 55 and 64 would consume 472.86 more ounces than households in the base category.

Households where the household head is employed part-time would consume 136.26 ounces less of conventional white milk than households where the household head is not employed for full pay. Households with a household head who has full time employment consume 283.05 less ounces of conventional white milk than the base category. Households with a household head who is college educated would consume 195.58 less ounces than households where the household head has less than a high school diploma. Post-college educated household heads would consume even less conventional white milk, consuming approximately 250 ounces less than the base category.

Households where the household head identifies as black would consume 1,109.63 ounces less than households where the household head identifies as white. Asian households would consume 671.62 ounces less than white households. Households that identify as other also would consume less with milk than the base category, consuming 395.63 ounces less than white households. Household heads who are of Hispanic-origin have households that would consume less conventional white milk than non-Hispanic households.

The households in the Mountain region are the only households that would consume less conventional white milk than households in the Pacific region. These households would consume 142.78 ounces less of conventional white milk than households in the Pacific region. Households in New England would consume 337.97

ounces more, households in the Middle Atlantic region would consume 176.24 ounces more, households in the East North Central region would consume 195.50 ounces more, households in the West North Central region would consume 577.94 ounces more, households in the South Atlantic region would consume 314.94 ounces more, and households in the East South Central region would consume 278.56 ounces more than the base category.

Households that have children present in the home would consume more conventional white milk than households where there are not any children. Households with children under 6 years old, between the ages of 6 and 12, and between the ages of 13 and 17 would consume 2,414.52 ounces more than households with no children. These households with children would consume at least 908.64 ounces more than the base category of households without children.

Households with a female household head only would consume 931.55 ounces less of conventional white milk that households with both a male and female household head. Households with a male household head would consume 841.81 ounces less than the base category.

For lactose-free milk, households where the household head is employed full time would consume 12.67 ounces less lactose-free milk than households where the household head is not employed for full pay. Households where the household head has a college education would consume 14.91 more ounces that households where the household head has less than a high school education. Post-college educated households would consume 22.57 more ounces of lactose-free milk than the base category of households.

Households that identified as black would consume the most lactose-free milk, consuming 44.27 more ounces than households that identified as white. Asian households would consume 28.74 more ounces than white households. Households that identify as other would consume 21.98 ounces more of lactose-free milk than households that identify as white. Hispanic-origin households would consume 28.25 ounces more than non-Hispanic households.

Households in the East North Central region and West North Central region would consume 36.39 ounces and 46.51 ounces less, respectively, than households in the Pacific region. Households in the South Atlantic region would consume 13.20 ounces less of lactose-free milk than the base category. Households in the East South Central region and in the West South Central region would consume 50.05 and 40.50 ounces less respectively, than households in the Pacific region. Finally, households in the Mountain region would consume 26.06 ounces less of lactose-free milk than households in the Pacific region.

Households with children less than 6 years old, and households with children less than 6 years old and children between the ages of 6 and 12 would consume the most lactose-free milk. These households would consume 15.55 ounces and 13.60 ounces more than households with no children.

Households with a male household head only would consume 17.61 ounces less of lactose-free milk than the base category of households with both a male and female household head.

Table 12 Unconditional and Conditional Own-price, Cross-price and Income Elasticities Soymilk, Almond Milk, Conventional White Milk, and Lactose-free Milk

Unconditional Own-Price, Cross-Price and Income Elasticities

	Soymilk	Almond Milk	Conventional White Milk	Lactose-free Milk	Income
Soymilk	**-3.37**	**-2.05**	**0.98**	**-1.07**	**0.17**
Almond Milk	**-1.18**	**-2.72**	**1.01**	**-2.30**	**0.26**
Conventional White Milk	0.03	**-0.14**	**-0.97**	0.08	**0.01**
Lactose-free Milk	-0.25	**-1.19**	**1.42**	**-2.85**	**0.23**

Conditional Own-Price, Cross-Price and Income Elasticities

	Soymilk	Almond Milk	Conventional White Milk	Lactose-free Milk	Income
Soymilk	**-0.67**	**-0.41**	**0.19**	**-0.21**	**0.03**
Almond Milk	**-0.24**	**-0.55**	**0.20**	**-0.46**	**0.05**
Conventional White Milk	0.02	**-0.10**	**-0.69**	0.06	**0.01**
Lactose-free Milk	-0.07	**-0.34**	**0.41**	**-0.49**	**0.07**

Numbers in bold font are statistically significant at *p*-value 0.05

Table 12 displays the conditional and unconditional own-price, cross-price and income elasticities for the four beverage categories. The conditional own-price elasticities of demand for soymilk, almond milk, lactose-free milk and conventional white milk are, -0.67, -0.55, -0.49 and -0.69 respectively. These elasticities indicate that consumers of these beverages are relatively insensitive to own-price changes, or those who purchase these beverages are loyal to purchasing the respective beverages. The unconditional own-price elasticities of demand for soymilk, almond milk, lactose-free milk and conventional white milk are, -3.37, -2.72, -2.85 and -0.97 respectively. These estimates are consistently larger than the conditional elasticities. This result shows a higher own-price and income response, as well as, more substitutability between beverages when all households that would potentially buy soymilk, almond milk, lactose-free milk and conventional white milk are taken into account.

Nine of the twelve conditional cross-price elasticities are significant. In the soymilk equation, almond milk and lactose-free milk are gross complements with conditional cross-price elasticities of -0.41 and -0.21 respectively. Conventional white milk is a gross substitute for soymilk with a conditional cross-price elasticity of 0.19. In the almond milk equation, soymilk and lactose-free milk are gross complements with conditional cross-price elasticities of -0.24 and -0.46 respectively. Conventional white milk is a gross substitute for almond milk with a conditional cross-price elasticity of 0.20. Only the almond milk conditional cross-price elasticity is significant in the conventional white milk equation. The conditional cross-price elasticity is -0.10 meaning that conventional white milk and almond milk are gross complements in consumption. In the lactose-free milk equation, almond milk is a gross complement with a conditional cross-price elasticity of -0.34. Conventional white milk is a gross substitute for lactose-free milk with a conditional cross-price elasticity of 0.41. The conditional income elasticities are 0.03 for soymilk, 0.05 for almond milk, 0.01 for conventional white milk and 0.07 for lactose-free milk. These results indicate that all four beverages are normal goods.

Like the own-price elasticities, the unconditional cross-price elasticities are larger than the conditional cross-price elasticities. The unconditional cross-price elasticity of soymilk with respect to almond milk is -2.05 meaning they are gross complements consumption. The unconditional cross-price elasticity of soymilk with respect to lactose-free milk is -1.07, again demonstrating that the two beverages are gross complements. Both soymilk and lactose-free are gross complements in consumption with respect to almond milk with unconditional cross-price elasticities of -1.18 and -2.30 respectively. The unconditional cross-price elasticity of conventional white milk with respect to almond milk is -0.14 meaning these two beverages are gross complements in consumption. The unconditional cross-price elasticity of lactose-free milk with respect to almond milk is -1.19 demonstrating that the beverages are gross complements in consumption. The unconditional cross-price elasticity of lactose-free milk with respect to conventional white milk is 1.42 demonstrating that these two beverages are gross substitutes. The unconditional income elasticities for soymilk, almond milk, conventional white milk and lactose-free milk are 0.17, 0.26, 0.01 and 0.23 respectively. These income elasticities indicate that all four beverage categories are normal goods.

7. CONCLUSIONS

The dairy alternative beverage market is changing the way Americans look at conventional milk products, which have long been a staple in the American diet. Using household-level purchase data of almond milk, soymilk, conventional white milk, and lactose-free milk, along with several associated demographic variables, we estimated the household level consumer demand for each of the beverage categories. Due to the censored nature of the data set, we used the Tobit model (Tobin, 1958) to estimate the conditional and unconditional factors affecting demand for these conventional dairy and dairy alternative beverages. This procedure allowed us to not only shed light on the probability of consumption but also allowed us to determine whether the beverages are gross substitutes or gross complements in consumption.

Employment status, education, race, Hispanic-origin, region, and gender are all significant factors determining demand for almond milk. The conditional own-price elasticity of demand is estimated to be -0.55. The cross-price elasticities between almond milk and soymilk, and almond milk and lactose-free milk reveal that these beverage products are gross complements in consumption. Consumers who purchase almond milk see conventional white milk as a gross substitute, although consumers who purchase conventional white milk view almond milk as a gross complement. Employment status, education, race, Hispanic-origin, region, age and presence of children, and gender are significant factors affecting the demand for conventional white milk.

The results found in this thesis complement the results from Dharmasena and Capps (2014) except, they found that the age and presence of children was a significant demographic variable affecting the demand for soymilk. Our results indicate that the age of the household head, employment status, education, race, Hispanic-origin, region, and gender are the significant demographic determinants of demand for soymilk. The cross-price elasticities found both in this thesis and Dharmasena and Capps (2014) indicate that conventional white milk is gross substitute for soymilk. The cross-price elasticities also indicate that soymilk is a gross complement to both almond milk and lactose-free milk.

Employment status, education, race, Hispanic-origin, and region significantly affect demand for lactose-free milk. Lactose-free milk demonstrates gross complementary behavior with almond milk with negative cross-price elasticity. As to be expected, lactose-free milk is a gross substitute to conventional white milk.

As far as limitations of this study, it must be noted that these results only indicate at home consumption/purchase of beverage products. The data used did not allow us to capture away-from home consumption and purchasing behavior. Although, these products are not prominently purchased for away-from home consumption.

These elasticity estimates and demographic information could be used by beverage manufacturers and marketers to strategically position dairy alternative beverages in the conventional dairy marketplace. This information can also be used by dairy marketers to stay competitive in this ever-changing marketplace. For example, the results from this study show that almond milk is a gross complement to conventional white milk. This result indicates that consumers are purchasing both conventional white milk and almond milk. Therefore, consumers are potentially replacing conventional white milk with almond milk for some uses. For example, consumers could use almond milk instead of conventional white milk in smoothies, cereal, etc. Marketers can use the demographic information gleaned from this study to target consumers who are already purchasing their products (the conditional sample) or all consumers (the unconditional sample) and therefore stay competitive in this diverse market.

Furthermore, government policy makers can use these elasticity estimates and demographic profiles of dairy alternative beverage consumers to design appropriate policies for U.S. dairy as well as dairy alternative industries.

Potential future research would be to use the elasticity estimates to simulate effects of dairy farmer welfare in the United States, in light of the rising competition from dairy alternative beverage industry.

REFERENCES

Alviola, P.A., and O.Capps, Jr. 2010. "Household Demand Analysis of Organic and Conventional Fluid Milk in the United States based on the 2004 Nielsen Homescan Panel." *Agribusiness* 26(3): 369-388.

Berry, S., Levinsohn, J., and A. Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63(4): 841-890.

Capps, Jr. O., R. Tsai, R. Kirby, and G. Williams. 1994. "A Comparison of Demand for Meat Products in the Pacific Rim Region." *Journal of Agricultural and Applied Economics* 19(1): 210-224.

Choi, H.J., and M. Wohlgenant. 2012. "Demand Analysis of Fluid Milk with Different Attributes." Selected Paper for Agricultural & Applied Economics Association's Annual Meeting. August 2012.

Davis, Christopher G., Dong, D., Blayney, D., Yen, S., and R. Stillman. 2012. "U.S. Fluid Milk Demand: A Disaggregated Approach." *International Food and Agribusiness Management Review* 15(1): 25-50.

Deaton, A., and J. Muelbauer. 1980. "An Almost Ideal Demand System." *American Economic Review* 70(3): 312-326

Dharmasena, S. and O.Capps Jr. 2014. "Unraveling Demand for Dairy-Alternative Beverages in the United States: The Case of Soymilk." *Agricultural and Resource Economics Review* 43(1): 140-157.

Glaser, L.K., and G.D. Thompson. 2000. "Demand for Organic and Conventional Beverage Milk." Paper presented at Western Agricultural Economics Association Annual Meetings, June 29-July 1. Vancouver, British Colombia, Canada.

Gould, B.W. 1996. "Factors Affecting U.S. Demand for Reduced-Fat Milk." *Journal of Agricultural and Resource Economics* 21: 68-81.

Gould, B.W., T.L. Cox, and F. Perali. 1990. "The Demand for Fluid Milk Products in the U.S.: A Demand Systems Approach." *Western Journal of Agricultural Economics* 15(1): 1-12.

Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47: 153-161.

Innova Market Insights. 2013. “Dairy Alternative Drinks Move Beyond Soy.” Food BusinessNews.net, accessed September 3, 2014.

Kaiser, H.M., and J.C. Reberte. 1996. "Impact of Generic Advertising of Fluid Milk on Demand for Whole, Low-fat, and Skim Milks." *Journal of Dairy Science* 79:2284-2291.

Kennedy, P. 2003. “Limited Dependent Variables.” *A Guide to Econometrics* 5th Edition, MIT Press, Cambridge, Massachusetts.

Kyureghian, G., O. Capps, Jr., and R. Nayga. 2011. “A Missing Variable Imputation Methodology with an Empirical Application.” *Advances in Econometrics* 27A:313-337.

Lopez, E., and R.A. Lopez. 2009. “Demand for Differentiated Milk Products: Implications for Price Competition. *Agribusiness* 25: 453-465.

McCormack, Ryan, 2016. “Dairy Product Production in the US.” IBISWorld, http://clients1.ibisworld.com/reports/us/industry/default.aspx?entid=244, accessed January 9, 2016.

McDonald, J. F., and R. A. Moffitt. 1980. “The Uses of Tobit Analysis.” *The Review of Economics and Statistics* 62(2): 318-321.

Neville, Antal. 2015. “Soy and Almond Milk Production in the US.” IBISWorld, http://clients1.ibisworld.com/reports/us/industry/default.aspx?entid=4195, accessed April 1, 2014.

Physicians Committee for Responsible Medicine. “What is Lactose Intolerance?” http://www.pcrm.org/health/diets/vegdiets/what-is-lactose-intolerance, accessed on May 30, 2016.

Statista. 2016. “Lactose-free Product Sales in the United States from 2012 to 2017, by Category (in million U.S. dollars).” http://www.statista.com/statistics/251134/us-lactose-free-product-sales-by-category/, accessed May 30, 2016.

Tobin, J. 1958. “Estimation of Relationships for Limited Dependent Variables.” *Econometrica* 26(1): 24-36.

USDA-ERS Economic Research Service (ERS). 2015. Dairy Data, http://www.ers.usda.gov/data-products/dairy-data.aspx, accessed on February 20, 2016.

USDA-ERS Economic Research Service (ERS). 2015. Food Availability Spreadsheets, United States Department of Agriculture-Economic Research Service, Washington, DC.

Van Allen, Jennifer, 2014. “Got Milk? From a Cow or a Plant?” The Washington Post, https://www.washingtonpost.com/news/to-your-health/wp/2014/06/10/got-milk-from-a-cow-or-a-plant/, accessed on September 3, 2014.

Zheng, Y. 2011. “Valuation of Country of Origins of Organic Processed Food: A Comparative Study of Consumer Demand for Soymilk in The United States and China.” A Thesis at Kansas State University.

APPENDIX A

DERIVATION OF EQUATION (20)

(21) $\left(\frac{\partial E(y)}{\partial X}\right) = F(z)\left(\frac{\partial E(y^*)}{\partial X}\right) + E(y^*)\left(\frac{\partial F(z)}{\partial X}\right)$

(22) $\left(\frac{\partial E(y)}{\partial X}\right) - F(z)\left(\frac{\partial E(y^*)}{\partial X}\right) = E(y^*)\left(\frac{\partial F(z)}{\partial X}\right)$

(23) $\left(\frac{\partial F(z)}{\partial X}\right) = \frac{1}{E(y^*)}\left(\frac{\partial E(y)}{\partial X} - F(z)\frac{\partial E(y^*)}{\partial X}\right)$

Where the normalized index value z is shown as $z = \frac{X\beta}{\sigma}$, σ is the estimated standard error of the tobit regression, $F(z)$ represents the cumulative distribution function (CDF) and $f(z)$ represents the probability density function (pdf) associated with the normalized index value, *z*. $E(y)$ represents the unconditional expected value (observations that are at or above the limit) and $E(y^*)$ represents the conditional expected value (observations that are above the limit).

APPENDIX B

DEMOGRAPHIC VARIABLES

Table 13 SAS Code for each Demographic Variable

Variable	Code
Age of household head 25-29	agehh2529
Age of household head 30-34	agehh3034
Age of household head 35-44	agehh3544
Age of household head 45-54	agehh4554
Age of household head 55-64	agehh5564
Age of household head >65	agehhgt64
Employment status part-time	emphhpt
Employment status full-time	emphhft
Education: high school	eduhhhs
Education: undergraduate	eduhhu
Education: post-college	eduhhpc
Black	black
Asian	asian
Other	other
Hispanic	hisp_yes
New England	newengland
Middle Atlantic	middleatlantic
East North Central	eastnorthcentral
West North Central	westnorthcentral
South Atlantic	southatlantic
East South Central	eastsouthcentral
West South Central	westsouthcentral
Mountain	mountain
Children less than 6 years	aclt6_only
Children 6-12 years	ac6_12only
Children 13-17 years	ac13_17only
Children < 6 & 6-12 years	aclt6_6_12only
Children <6 & 13-17years	aclt6_13_17only
Children 6-12 & 13-17 years	ac6_12and13_17only
Children <6 & 6-12 & 13-17	aclt6_6_12and13_17only
Female head only	fhonly
Male head only	mhonly

APPENDIX C

SAS CODE

1. Soy Milk

```
*price imputation auxilliary regression for Soymilk;
proc reg data=thesis.CD_one;
model newprice_soy=L_hinc hsize NewEngland MiddleAtlantic
EastNorthCentral WestNorthCentral SouthAtlantic
EastSouthCentral WestSouthCentral Mountain;
run;

data thesis.CD_one; set thesis.CD_one;
if (total_soy_oz ne 0 or total_soy_exp ne 0) then
newprice_soy=total_soy_exp/total_soy_oz;
if (total_soy_oz=0 or total_soy_exp=0)
then newprice_soy=0.05646 + 0.00006875*L_hinc -
0.00054610*hsize + 0.00196*NewEngland +
0.00189*MiddleAtlantic - 0.00031621*EastNorthCentral -
0.00083359*WestNorthCentral - 0.00160*SouthAtlantic -
0.00141*EastSouthCentral - 0.00318*WestSouthCentral -
0.00120*Mountain;
run;

*Following is the tobit model for Soymilk;
Proc QLIM data=thesis.CD_one;
model Q_soy=L_P_soy L_P_almond L_P_W L_P_LF L_hinc
agehh2529 agehh3034 agehh3544 agehh4554 agehh5564 agehhgt64
emphhpt emphhft eduhhhs eduhhu eduhhpc
NewEngland MiddleAtlantic EastNorthCentral WestNorthCentral
SouthAtlantic EastSouthCentral WestSouthCentral Mountain
black asian other hisp_yes aclt6_only ac6_12only
ac13_17only aclt6_6_12only aclt6_13_17only
ac6_12and13_17only aclt6_6_12and13_17 fhonly mhonly;
endogenous Q_soy ~ censored(lowerbound=0);
nloptions maxiter=500;
output out=thesis.tobit_soy conditional expected marginal
xbeta residual;
run;
```

2. Almond milk

```
*price imputation auxilliary regression for Almond Milk;
proc reg data=thesis.CD_one;
model newprice_almond=L_hinc hsize NewEngland
MiddleAtlantic EastNorthCentral WestNorthCentral
SouthAtlantic EastSouthCentral WestSouthCentral Mountain;
run;

data thesis.CD_one; set thesis.CD_one;
if (total_almond_oz ne 0 or total_almond_exp ne 0) then
newprice_almond=total_almond_exp/total_almond_oz;
if (total_almond_oz=0 or total_almond_exp=0)
then newprice_almond=0.05439 + 0.00055220*L_hinc -
0.00059414*hsize - 0.00114*NewEngland -
0.00012344*MiddleAtlantic - 0.00120*EastNorthCentral -
0.00073378*WestNorthCentral - 0.00374*SouthAtlantic -
0.00459*EastSouthCentral - 0.00539*WestSouthCentral -
0.00205*Mountain;
run;

*Following is the tobit model for Almond Milk;
Proc QLIM data=thesis.CD_one;
model Q_almond=L_P_soy L_P_almond L_P_W L_P_LF L_hinc
agehh2529 agehh3034 agehh3544 agehh4554 agehh5564 agehhgt64
emphhpt emphhft eduhhhs eduhhu eduhhpc
NewEngland MiddleAtlantic EastNorthCentral WestNorthCentral
SouthAtlantic EastSouthCentral WestSouthCentral Mountain
black asian other hisp_yes aclt6_only ac6_12only
ac13_17only aclt6_6_12only aclt6_13_17only
ac6_12and13_17only aclt6_6_12and13_17 fhonly mhonly;
endogenous Q_almond ~ censored(lowerbound=0);
nloptions maxiter=500;
output out=thesis.tobit_almond conditional expected
marginal xbeta residual;
run;
```

3. Conventional white milk

```
*price imputation auxilliary regression for conventional
white milk;
proc reg data=thesis.CD_one;
```

```
model newprice_W= L_hinc hsize NewEngland MiddleAtlantic
EastNorthCentral WestNorthCentral SouthAtlantic
EastSouthCentral WestSouthCentral Mountain;
run;

data thesis.CD_one; set thesis.CD_one;
if (total_W_oz ne 0 or total_W_exp ne 0) then
newprice_W=total_W_exp/total_W_oz;
if (total_W_oz=0 or total_W_exp=0)
then newprice_W=0.03092 + 0.00095979*L_hinc - 0.00177*hsize
+ 0.00203*NewEngland + 0.00126*MiddleAtlantic -
0.00473*EastNorthCentral - 0.00154*WestNorthCentral +
0.00355*SouthAtlantic + 0.00072782*EastSouthCentral +
0.00029510*WestSouthCentral - 0.00531*Mountain;
run;

*Following is the tobit model for conventional white milk;
Proc QLIM data=thesis.CD_one;
model Q_W=L_P_soy L_P_almond L_P_W L_P_LF L_hinc agehh2529
agehh3034 agehh3544 agehh4554 agehh5564 agehhgt64 emphhpt
emphhft eduhhhs eduhhu eduhhpc
NewEngland MiddleAtlantic EastNorthCentral WestNorthCentral
SouthAtlantic EastSouthCentral WestSouthCentral Mountain
black asian other hisp_yes aclt6_only ac6_12only
ac13_17only aclt6_6_12only aclt6_13_17only
ac6_12and13_17only aclt6_6_12and13_17 fhonly mhonly;
endogenous Q_W ~ censored(lowerbound=0);
nloptions maxiter=500;
output out=thesis.tobit_W conditional expected marginal
xbeta residual;
run;
```

4. Lactose-free milk

```
*price imputation auxilliary regression for lactose free
milk;
proc reg data=thesis.CD_one;
model newprice_LF= L_hinc hsize NewEngland MiddleAtlantic
EastNorthCentral WestNorthCentral SouthAtlantic
EastSouthCentral WestSouthCentral Mountain;
run;

data thesis.CD_one; set thesis.CD_one;
```

```
if (total_LF_oz ne 0 or total_LF_exp ne 0) then
newprice_LF=total_LF_exp/total_LF_oz;
if (total_LF_oz=0 or total_LF_exp=0)
then newprice_LF=0.05829 + 0.00120*L_hinc -
0.00049151*hsize - 0.00595*NewEngland -
0.00166*MiddleAtlantic - 0.00804*EastNorthCentral -
0.00624*WestNorthCentral - 0.00594*SouthAtlantic -
0.00773*EastSouthCentral - 0.01197*WestSouthCentral -
0.00523*Mountain;
run;

*Following is the tobit model for lactose free milk;
Proc QLIM data=thesis.CD_one;
model Q_LF=L_P_soy L_P_almond L_P_W L_P_LF L_hinc agehh2529
agehh3034 agehh3544 agehh4554 agehh5564 agehhgt64 emphhpt
emphhft eduhhhs eduhhu eduhhpc
NewEngland MiddleAtlantic EastNorthCentral WestNorthCentral
SouthAtlantic EastSouthCentral WestSouthCentral Mountain
black asian other hisp_yes aclt6_only ac6_12only
ac13_17only aclt6_6_12only aclt6_13_17only
ac6_12and13_17only aclt6_6_12and13_17 fhonly mhonly;
endogenous Q_LF ~ censored(lowerbound=0);
nloptions maxiter=500;
output out=thesis.tobit_LF conditional expected marginal
xbeta residual;
run;
```